Lincoln University
Te Whare Wānaka o Aoraki
CHRISTCHURCH•NEW ZEALAND

Lincoln University Digital Dissertation

Copyright Statement

Assessing honeybees' use of *Phacelia tanacetifolia*, a key component of agri-environment schemes, using molecular and field techniques

---

A Dissertation

submitted in partial fulfilment

of the requirements for the Degree of

Honours of Science

at

Lincoln University

by

Rowan Sprague

Lincoln University

2015

Abstract of a Dissertation submitted in partial fulfilment of the requirements for the Degree of Honours of Science

Assessing honeybees' use of *Phacelia tanacetifolia*, a key component of agri-environment schemes, using molecular and field techniques

by

Rowan Sprague

About 70% of crops worldwide benefit directly or indirectly from animal pollination, with insects contributing the most to pollination. Of these, honeybees (*Apis mellifera*) are used the most by farmers for pollination services because of their abundance, generalist floral preferences, and ease of management. Increased dependence on honeybees coupled with their recent population decline has caused the European Commission to enact programs aimed at improving pollinator fitness and efficacy, called Agri-Environment Schemes (AES). While these AESs include specific guidelines for farmers and landowners to follow, scientific assessment of these AESs is still lacking. This project seeks to address this issue by assessing how honeybees use the pollen of *Phacelia tanacetifolia*, a wildflower commonly used in AESs, in an agricultural landscape.

In order to identify accurately and efficiently the plant species of pollen that honeybees collect, this project researched existing pollen identification methods and examined Next Generation Sequencing (NGS) of pollen DNA as an alternative method. While many samples were able to be sequenced and identified to species level using NGS, it was found to be unsuitable for small-scale pollen identification projects due to its expense.

Based on the pollen sequencing results and observations in the field, this study found that honeybees did not use *P. tanacetifolia* much for pollen, although they did forage on *P. tanacetifolia* for nectar. In the wider context, the overall results of this study indicate that honeybees may not prefer the flowers provided in AESs for pollen depending on the other available floral resources in the surrounding area. Since AES measures are aimed at improving pollinator health, scientists and policy makers should ensure that the floral resources used in AESs are preferred by the target pollinator group.

Keywords: pollen identification, metabarcoding, Next Generation Sequencing, *Apis mellifera,* pollen preference, honeybee foraging behaviour, agroecosystems

Acknowledgements

This project could not have been completed without the help of many people. First, I would like to thank my supervisors, Steve Wratten, Stephane Boyer, and Rob Cruickshank, who helped me with the experimental design, guidance in laboratory and field techniques, data processing, and the editing of this document and many others. I would also like to thank Steve Wratten for his funding assistance for the Next Generation Sequencing method. Dan Arthur, a beekeeper in Lincoln, lent his honeybee hives for this research, assisted in installing the pollen traps, and allowed me to plant *Phacelia tanacetifolia* on his land near the beehives; a big thank you to him for his help. I am also very grateful to Andrew Frapwell who lent two of his hives to this research for collecting pollen in 2013-2014.

I had a lot of help from two summer scholars in this project. Jessica Watson, a summer scholar from November 2013-February 2014, helped collect the pollen samples used in the Next Generation Sequencing method, and Georgia Stevenson, a summer scholar from November 2014-February 2015, collected the observational data on honeybees' use of *P. tanacetifolia* at different times of day. Elizabeth Permina, a bioinformatician at NZGL, de-multiplexed and de-noised the sequences from the Next Generation Sequencing. Brian Kwan, Myles MacKintosh, Fariba Nourozi, and Norma Merrick helped me by ordering laboratory supplies, teaching me laboratory protocols, Sanger capillary sequencing pollen samples, and always answering my questions. Alan Marshall, Dave Jack, and Dan Stewart from the Field Service Centre provided me with vans for transporting *P. tanacetifolia* to my field site and cultivated my field site. Brent Richards and Leona from the Lincoln University Nursery also assisted me by provided seed trays and soil and also watering my plants while they were in the greenhouses. Thank you as well to Jennifer Bufford, Maud Bernard-Verdier, Will Godsoe, Ian Dickie, Maki Ikegami, and Tyler Brummer for their help with statistical methods and R coding; I learned a great deal from them about data analysis and good coding habits.

I would also like to thank Fulbright New Zealand for their funding support which enabled me to come to New Zealand and pursue my research aspirations. I am also grateful for Fulbright New Zealand's cultural and social support in welcoming me to New Zealand and helping me acclimate to living here. Similarly, thank you to the Bio-Protection Research Centre for their institutional support as well as friendly and collaborative environment; morning tea breaks with people from BPRC would almost always end in laughs, new ideas, or both.

Finally, I could not have done this project without the endless encouragement from my family, my partner Johnathon Ridden, and my friends. They helped with editing, and more importantly, they believed in me the whole time and helped me keep positive. I am forever grateful for their support.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background and significance

As much as 70% of crops worldwide benefit from indirect or direct animal pollination (Klein et al. 2007), with insects contributing the most to pollination. Of these, honeybees (*Apis mellifera*) are used most frequently by farmers for pollination services because of their abundance, generalist floral preferences, and ease of management and transportability (Tautz 2008; Aizen & Harder 2009; Potts et al. 2010a). Although the significance of native or unmanaged insects in crop pollination has been increasingly recognised (Winfree et al. 2008; Rader et al. 2009; Woodcock et al. 2013), reliance on honeybees has nevertheless increased in recent decades (Aizen et al. 2008; Breeze et al. 2011).

While demand is surging, honeybee populations are declining in many countries that are heavily reliant on honeybee services, such as the United States, Canada, the UK, and Germany (Potts et al. 2010a, 2010b; van der Zee et al. 2012; vanEngelsdorf et al. 2012). No single cause of this decline has been identified; rather, several factors have been found, notably varroa mites (*Varroa destructor*) (Sammataro et al. 2000; Shen et al. 2005), pathogens such as American Foulbrood (caused by the bacteria *Paenibacillus spp.*) (Genersch 2010), viruses caused by spores such as *Nosema* (Potts et al. 2010a; Pettis et al. 2013), loss of biodiversity and floral resources in agricultural systems (Kremen et al. 2007; Klein et al. 2007; Potts et al. 2010a), and the use of neonicotinoid pesticides (vanEngelsdorp & Meixner 2010; Potts et al. 2010a; Goulson 2013).

In response to this decline, the European Commission has recommended programmes aimed at improving pollinator fitness and efficacy as part of the wider fiscal policy of Agri-Environment Schemes (AESs). Originating from the 1980s to protect biodiversity and important cultural areas in England, these AESs inform farmers and landowners on methods to manage land sustainably and promote ecosystem services, or nature's services (Natural England 2012). Included in this policy are methods that claim to benefit pollinators, such as bees and butterflies (Natural England 2013a & 2013b), although no supporting scientific evidence is offered. AESs have evolved over time to incorporate more accessible ways for farmers to become involved and to include more management recommendations to benefit various ecosystem services. By 2009, about 66% of England's agricultural land was managed as part of AES agreements (Natural England 2009). Between 2007 and 2013, the EU spent 20 billion euros on AES, or 22% of the total amount spent on rural development in that region. In spite of AES being increasingly engrained in policy and costing

governments billions of euros, scientific assessments of these schemes are still lacking (Kleijn & Sutherland 2003; Kleijn et al. 2006; Whittingham 2007).

Other countries are also attempting to address the issue of pollinator decline. The US Federal government has also recently proposed the Pollinator Health Strategy 2015, which seeks to improve the health of honeybees and native pollinators, and restore land to be used as pollinator habitat. Different from the AES Environmental Stewardship handbooks, this report acknowledged the elements in its strategy which will to be researched before the government can recommend more successful policies (Vilsack & McCarthy 2015). However, both this policy report and the AES Handbooks lacked constructive and scientific ways to record and assess the results of the landscape enhancements made. Studies evaluating how insect pollinators use these floral resources and the extent to which they benefit from them are needed to assess how useful AESs really are for insect pollinators.

1.2 Pollen identification

When assessing the effectiveness of AESs for insect pollinators, it is very useful to know what plants the insects forage for and how much, and how often, they use the floral resources provided for them. One way to do this is to analyse the pollen collected by these pollinators (Kearns & Inouye 1993). When honeybees collect pollen from flowers, they group the pollen together to form pellets which they carry on their hind legs in specialised structures called corbiculae (Keller et al. 2005a). These pollen pellets can be gently taken off honeybees and analysed for botanical identification. A range of pollen identification methods have been used to determine from which plant species, genera, or families the honeybees are foraging and depending on the method, relative plant species or genera abundance in the collected pollen may also be determined (Dimou & Thrasyvoulou 2007).

Pollen identification methods vary greatly in terms of time and cost needed to complete them and also in the accuracy of the data collected. Some methods based on visual identification require a high level of expertise and require a reference library of all the possible species that could be found in the pollen samples collected from the insects. Six pollen identification methods have been identified in the literature: 1) Visual categorization of pollen pellets using the naked eye; 2) Microscopy of individual pollen grains; 3) Flow cytometry of pollen pellets; 4) Infrared spectroscopy of pollen grains; 5) Computer image recognition of individual pollen grains; and 6) Sequencing of DNA in pollen grain cells. These methods will be described below and discussed in terms of their efficacy for use in contributing to evaluate the effectiveness of AESs.

1) *Visual categorization* involves grouping pollen pellets collected from honeybees into colour groups (e.g., Arita et al. 1989) and then taking a small subsample of the pollen pellets in each colour group

and using microscopy of pollen grains to check the plant species. While this method can potentially be used to identify plant species based on distinct pollen colours (Stimec et al. 1997), it can also be inaccurate for many plant species. Pollen pellets categorized in one colour group can contain pollens of several different plant species despite colours appearing the same to the naked eye (Almeida et al. 2005). Pollen from one particular plant species can also display different colours; for example pollen in the genus *Myrcia* can be polychromatic (Barth et al. 2009). Rain and sun exposure can alter the colour of pollen while it is still on the anthers, as can the storage conditions under which it is kept (Barth et al. 2009; Stanley & Linskens 1974). For these reasons, visual categorization is no longer commonly used for precise identification of plant species from pollen, especially because of advances in other methods.

2) *Microscopy* is the conventional method used to identify pollen. It involves taking a small sample of pollen grains from a pollen pellet and treating them with acetic and sulphuric acid, followed by staining (e.g., Barth et al. 2010). Under the microscope, the pollen grains can be identified to species or genera by grain shape, surface texture, and internal cellular structure (Erdtman 1943; Moore et al. 1991). While this method is generally accurate to genus level, it cannot guarantee accurate identification to species (Almeida et al. 2005; Khansari et al. 2012). It can also be time-consuming and requires the use of strong acids for staining, thus needing specialised laboratory facilities for safe use. More importantly, it requires a high level of taxonomic expertise in pollen identification and specific training of the operator.

3) *Flow cytometry* is a relatively new method for pollen identification, recently proposed by Kron et al. (2014). Pollen grains can be passed through the UV laser beam of the flow cytometer, which can measure the genome size in individual pollen grains. Although the machine must be calibrated to group the grains into separate categories, it can be accurate as genome size differs greatly between plant species (Leitch & Bennett 2007). However, this method has not yet been shown to work for a large mixed group of an unknown number of pollen types, as Kron et al. (2014) only compared two species.

Two preliminary but potentially promising methods also exist: 4) *Infrared spectroscopy* and 5) *Computer imaging software*. Gonzalez-Martin et al. (2007) used Near Infrared (NIR) spectroscopy to assess the chemical composition of honeybee pollen, but they did not test this method to identify plant species. Gottardini et al. (2007) used Fourier Transform Infrared (FT-IR) Spectroscopy to examine samples of airborne pollen, but they tested the identification of pollen for only two plant species. Both these methods would need to be further tested to assess their utility for identifying plant species from a mixed sample of pollen pellets. The computer imaging software is also a potential pollen identification method as it uses four different texture measurements of pollen grains

to categorize pollen samples (Marcos et al. 2014). The method would also need to be developed further, with more calibrations of pollen samples and public accessibility of the software for it to be used in identification of pollen from an unknown number of species.

6) Another method is *DNA Sanger sequencing of pollen* (Wilson et al. 2010; Galimberti et al. 2014). The nucleus of the pollen grain contains plant DNA, which can be extracted, amplified, and sequenced using standard molecular techniques. Standard sequencing works well on an individual pollen pellet basis since individual honeybees typically forage on only one plant species each day (Barth et al. 2009; Tautz 2008; Free et al. 1963). However, when analysing the pollen from a honeybee colony, with thousands of individual foraging honeybees, there can be a large number of plant species collected during a day. In this case, the method would require sequencing a great number of samples to guarantee the identification of all species present. Therefore, for studies looking at plant species foraging by whole honeybee colonies, this method would be very time and cost inefficient.

The current project investigates using Next Generation Sequencing (NGS) of DNA as an extension of the DNA Sequencing method for pollen identification and quantification. NGS allows researchers to obtain large numbers of sequences from mixed samples (Shendure & Hanlee 2008). Applying these methods to pollen pellets collected from beehives would allow pollination studies to determine the plant species visited by foragers, without knowing beforehand the number of plant species present in each sample. Additionally, since NGS has the potential to sequence all amplicons within a sample, the pollen pellets can be mixed together instead of being analysed one at a time. The present study evaluates the potential for NGS as an alternative pollen identification method by measuring how effectively the method can identify the presence and abundance of plant species in pollen samples collected from honeybee hives.

Pollen identification would enable the effectiveness of AES to be assessed by determining on which plants the insects are foraging and how they are using the floral resources provided in the AES. If the results from pollen identification indicate that insects are not using these floral resources, then the AES designs should be changed to increase their appeal to insect pollinators. Identification of insect-collected pollen would allow for the quantification of insect pollinator preferences. While the floral resources included in AES are claimed to benefit pollinators, pollen identification would assist in determining whether insects use these resources preferentially, or if they prefer other species not included in the AES. Thus, relative pollen preference, determined through pollen identification and quantification methods, would serve as a way to evaluate if, and to what extent, insect pollinators are using the supposedly beneficial plants in such schemes.

1.3 Pollen preference

Relative pollen preference in the case of honeybees means that the colony as a whole shows a preference for the pollen of a particular plant species relative to the surrounding flowering plants (Levin & Bohart 1955). While honeybee preference for flowers depends on many factors, including location of the plants in the landscape, morphology of the flowers, and flower colour and odour, honeybees tend to prefer the pollen of certain plant species over others. Boch (1982) found that colour of pollen can influence pollen preference, with yellow pollen being more preferred. However, colour preference for pollen has not been widely investigated in the literature. In controlled greenhouse experiments, Schmidt (1982) found that honeybees tended to prefer maple and almond pollen over the pollen of six other species. Olsen et al. (1979) trapped pollen from honeybee hives located near apple, blueberry, cucumber, and strawberry crops and compared pollen preference for the crops based on relative species representation in the pollen samples. The study found that apple and strawberry pollen were well represented in the samples, whereas cucumber pollen was not present (Olsen et al. 1979). In spite of these observations, relative pollen preference in relation to AESs has not been examined.

As mentioned before, studies assessing the effectiveness and ecological benefits of AESs are lacking (Kleijn & Southerland 2003; Kleijn et al. 2006; Whittingham 2007), despite the requirement for EU member states to include them in their rural development plans (European Commission 2015). The Environmental Stewardship Handbook, which describes the rules and guidelines for AESs in the UK, claims that margins of flowers sown around cultivated fields are designed to benefit birds and invertebrates, including insect pollinators (Natural England 2013a & 2013b). However, these claims were not backed up with scientific evidence and need to be evaluated.

There are a few scientific studies that explore use of AESs by honeybees and the potential ways in which they benefit from them. Couvillon et al. (2014) examined honeybee waggle dance patterns to determine where they forage in landscapes containing different types of AES as well as rural lands without any such stewardship measures. This study found that honeybees showed a significant foraging preference for Higher Level Stewardship sites (Couvillon et al. 2014). However, this study did not explore whether or not this preference was correlated to the floral resources available. Carvell et al. (2007) found that bumble bee species abundance and diversity increased in response to field margins planted with legume-based pollen and nectar seeds, but they concluded that a diverse mix of flowers should be planted to offer varying blooming durations and flower phenology. No studies looked at the foraging preference of honeybees, or lack thereof, for particular floral resources used in AESs.

1.3.1 *Phacelia tanacetifolia*, a key species in AES

Phacelia tanacetifolia is a wildflower that is commonly planted in the field margins of AESs as a floral resource. Native to California, this annual purple flower can germinate year-round and has a long flowering period of 6-8 weeks (Bowie et al. 1995; Williams & Christian 1991). It is a known high-quality honey plant (Crane et al. 1984) and insect attractant (Carreck & Williams 1997). Its pollen has a high protein content (Trees for Bees NZ 2014) and its nectar and pollen have been shown to improve the fitness of insects, including hoverflies (Hickman & Wratten 1996; Laubertie et al. 2012).

Phacelia tanacetifolia have several floral traits which are of interest for use in AESs. The flowers of *P. tanacetifolia* secrete the highest amount of nectar in 4- to 7-hr old flowers (Williams 1997). A study by Ekroos et al. (2008) found that insect pollinators benefitted from field margins, a type of AES measure, which contained nectar-rich flower species. While it is known that *P. tanacetifolia* has a high-quality nectar, it is not known how much nectar the *P. tanacetifolia* flowers produce. Thus, the nectar flow of *P. tanacetifolia* should be examined. Since *P. tanacetifolia* is a key component of AESs and has high-quality nectar and pollen, preference of honeybees for pollen of this species was studied as a way to assess part of AESs. Figure 1.3.1 shows *P. tanacetifolia* planted as a crop margin, a type of AES measure.

Figure 1.3.1 *Phacelia tanacetifolia* as it would look planted as a crop margin alongside an agricultural field.

1.3.2 Foraging behaviour of honeybees

Honeybees gather both nectar and pollen from flowers. Nectar brought back to the hive is eaten by the worker bees and also dehydrated and stored in the form of honey (Sammataro & Avitabile 2011; Tautz 2008). Pollen brought back to the hive is converted by the younger worker bees into a protein-rich jelly to be fed to the larvae and the queen. The higher the pollen quality, measured by protein, amino acid, antioxidant, lipid, and sugar contents, the fitter the honeybee larvae and queen will be, leading to an overall healthier honeybee colony (Di Pasquale et al. 2013; Keller et al. 2005a). Most foraging honeybees collect nectar and only a minority collect pollen as it is a more specialised task (Tautz 2008). Foraging honeybees will collect nectar as long as there are flowers available and the weather is suitable (Fewell & Winston 1995), but foraging honeybees will collect pollen only if the hive needs it (Camazine 1993; Dreller & Tarpy 2000). This usually occurs when the queen is laying many eggs a day, or when the hive is stocking up on stores for the winter (Keller et al. 2005b).

Honeybees were used as the study organism for this project not only because of their agricultural importance in providing pollination services, but also because of their distinct foraging behaviours. Individual honeybees collect pollen by gathering the pollen grains from the anthers of flowers and they use nectar to keep the pollen grains together. They carry the pollen in pollen baskets, or corbiculae, on their hind legs, forming pellets of pollen grains (Keller et al. 2005a). Honeybees demonstrate floral constancy, meaning that individual foraging bees will visit only one species of flower in any one day, sometimes even over several days (Free 1963). As a result of constancy of individual honeybees for one plant species, each pollen pellet is 95-99% of the time comprised of only one species (Tautz 2008; Newstrom-Lloyd et al. n.d.).

There are many factors that can influence on which species a honeybee will forage. Species of flowers can be chosen on the basis of their petal colour, as honeybees are shown to prefer yellow and blue (Goulson et al. 2007); their odour, as honeybee can be attracted to or repelled by odour (Ragubeer et al. 2013; Pernal & Currie 2002); ease of access to nectar and pollen, as honeybees will visit a species more if they can easily gather nectar and/or pollen from it (Pernal & Currie 2002); and density of blossoms, as honeybees prefer plants with a high density of flowers. Different species of flowers can also have varying diurnal patterns of nectar flow, with some species having only high nectar flow in the morning and some only in the afternoon. Desirability of species can also be complicated by where the flowers are located relative to the hive (Henry et al. 2012). If the flowers are farther away, or if they are obscured by landscape features, then the species may not be preferred by the colony.

1.4 Project aim and research questions

The overall aim of this project was to contribute to assessing the effectiveness of AESs for honeybees, first by examining pollen identification methods useful in determining on which plant species honeybee colonies forage and second by testing relative preference of honeybees for *P. tanacetifolia* pollen, a plant commonly found in AESs.

The following research questions were used to address the overall project aim:

Part One: Pollen Identification Methods

1. How do pollen identification techniques compare against each other in regards to time requirements, expense, and accuracy?

2. How well can NGS detect species presence and species abundance?

3. Compared to the pollen identification methods researched in the literature, how effective is NGS as a tool for assessments of honeybees' use of floral resources in AESs?

Part Two: Evaluating the Use of *P. tanacetifolia* by honeybees

1. To what extent do honeybees forage for pollen on *P. tanacetifolia* in an agricultural landscape?

2. What are the diurnal patterns of nectar flow in *P. tanacetifolia*?

Chapter 2

Methods

This section will be divided into two parts: one discussing methods for pollen identification and one examining honeybees' use of *Phacelia tanacetifolia* as a way to assess Agri-Environment Schemes (AESs).

2.1 Pollen identification methods

Six pollen identification methods were researched using the literature to compare their advantages and disadvantages as well as costs and time demands relative to each other. This comparison will answer the first research question of how pollen identification techniques compare against each other. Then the following methodology was used to address the second and third research questions on evaluating Next Generation Sequencing (NGS) as an alternative pollen identification method.

2.1.1 Field collection

Pollen was collected from two healthy honeybee colonies during the summer of 2013-2014 in the Canterbury province, New Zealand (latitude: -43.63788760, longitude: 172.53225660). Pollen traps were attached to the bottom of each hive and used to collect pollen from the honeybees. Pollen traps have two entrances which allow the foraging honeybees to enter and exit the hive freely. The main entrance/exit can be closed, forcing all of the foraging bees to enter and exit through a smaller entrance. This smaller entrance makes the honeybees entering the hive go through a fine mesh grid, effectively brushing off any pollen pellets on the hind legs of the foraging honeybees. These pollen pellets can be collected from the pollen traps via a sliding tray located under the mesh grid. Pollen traps are the recommended method for collecting pollen from a honeybee bee (Dimou & Thrasyvoulou 2007) and this method is used very commonly in the literature on pollen foraging behaviours of honeybees.

Pollen was collected on days when the temperature was above 18°C to ensure consistently high bee foraging activity. Collection times were 10:00-11:00 in the morning and 14:00-15:00 in the afternoon to compare plant species foraged for during these two periods. Pollen was collected for a total of 11 days. The pollen was collected by removing the bottom tray of the pollen trap and emptying all of the pollen pellets into a 25 mL tube. The collected pollen was then stored at -80°C for subsequent DNA analysis.

2.1.2 DNA extraction

To prepare the samples for DNA extraction, the pollen pellets were frozen with liquid nitrogen and then ground into a fine powder using a mortar and pestle. The mortar and pestle had been bleached overnight and oven-dried between each use to ensure no cross-contamination between samples. From each pollen powder sample, three subsamples (triplicates) were taken to evaluate replicability and accuracy of NGS results by cross-checking species presence/absence and abundance between subsamples. A GF-1 Nucleic Acid Extraction kit (Vivantis Technologies) was used to extract DNA from the pollen samples. Once obtained, the extracted DNA was stored at -20 °C.

2.1.3 Polymerase Chain Reactions (PCR)

The ITS (Internal Transcribed Spacer) gene region was selected for amplification of DNA extracts by PCR. ITS was selected as universal primers are available for this region and it is easy to amplify due to its high copy number in the genome; for this project the ITS primers ITS 5 was used for forward sequences and ITS 2 for reverse sequences (Baldwin 1992; Baldwin et al. 1995; Salariato et al. 2015). To amplify the DNA sequences properly for use in NGS, fusion primers were designed. These primers differ from standard primers in that they contain the following additions to the forward or reverse primer sequence: (1) adaptor sequences that are needed for 454 pyrosequencing and (2) Multiplex IDentifier (MID) tags, to identify each unique sample. Example forward and reverse fusion primer sequences are shown below in Figure 2.1.1; for a complete list of fusion primer sequences, see Appendix 1.

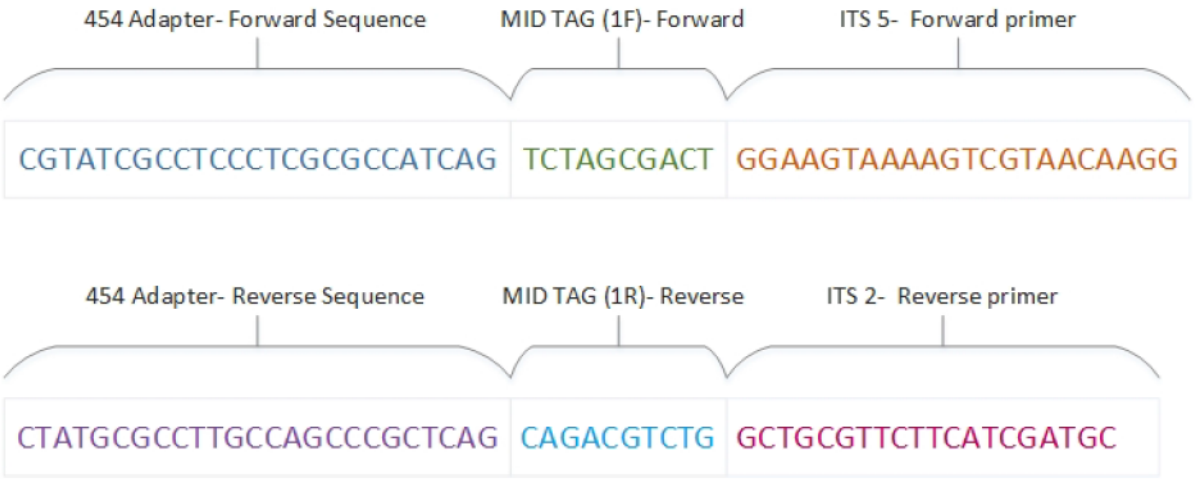

Figure 2.1.1 Example of forward and reverse fusion primers

Each PCR reaction contained a total of 10 µL, made up of 2.7 µL ultra-pure DNA-free water, 5 µL of Gotaq Green (Promega, Madison, WI, USA), 0.4 µL of forward and reverse fusion primers (10 µM), and 1.5 µL of DNA extract. Positive and negative controls were used for each PCR run to check for contamination. The PCR reaction conditions in the Thermocycler were optimised from Baldwin (1992) as follows: 97 °C for 1 minute for initial denaturation; 30 cycles at 97 oC for 1 minute, 53 °C for 1 minute, and 72 °C for 45 seconds; and finally 72 °C for 7 minutes for final extension. All PCR products were stored at 4 °C.

2.1.4 Gel extraction and purification

PCR products were run through a 1.5% agarose gel by electrophoresis and visualised with SYBR Safe DNA stain. Under blue light, all visible bands were cut out of the gel using a sterile scalpel and purified following the Zymo Gel Purification Kit instructions (Zymo Research). Purified PCR products were eluted in 10 µL of water and stored at -20°C.

2.1.5 DNA concentration normalisation

DNA concentrations were measured using a Qubit Assay kit dsDNA HS (0.2-100ng) and a Qubit fluorometer (Life Technologies, Auckland, New Zealand). PCR and gel extraction were repeated for any sample with DNA concentration below 1 µM. All samples were normalised to 2 ng/µL DNA concentration. If a sample was above 2 ng/µL then the appropriate amount of DNA-free water was added to dilute the concentration; if a sample was below 2 ng/µL but above 1 ng/µL, then more PCR product was added to reach 2 ng/µL. The samples were then pooled for sequencing.

2.1.6 454 sequencing and data processing

Pooled PCR products were sent to the AGTC laboratory at the University of Kentucky for 454 - pyrosequencing and run on 3 lanes of a GS-Junior Titanium system to test for potential systematic bias. The software PRINSEQ was used to trim and filter out poor quality sequences (Schmeider & Edwards 2011). The rest of the data processes for the NGS sequences were done using the software QIIME (Caporaso et al. 2010). All potential chimera sequences were removed, and the sequences were de-multiplexed (sorted) into samples and subsamples according to their MID tags and primers. Sequences were then de-noised to reduce the number of falsely identified sequences. For each sample, sequences were then collapsed to obtain a list of molecular taxonomic units (MOTUs) or representative sequences, which were queried against the database GenBank (Altschul et al. 1990). Only reads with more than 97% similarity with GenBank species (Caporaso et al. 2010) and which were known to be present in the geographical region were considered as valid matches. For each MOTU, the species match with the highest similarity percentage and lowest E-score was chosen. The

number of sequences which were collapsed into each MOTU was taken to represent relative abundance.

2.1.7 Data analysis

To answer the second research question of how well does NGS detect species presence and species abundance, the triplicates were compared within each sample. All of this analysis was done in R (R Core Team 2014). First the data were subset into individual samples containing the three triplicates. Then species absence data was removed when a species was absent from all triplicates in a sample to avoid skewing the analysis. Presence and abundance data were compared between the triplicates using generalised mixed effects linear models. The summary statistics were used to determine how well the triplicates correlated between each other, with a high correlation signifying a high level of accuracy and consistency of the NGS method.

The data generated from NGS was then used to investigate how species presence/absence and species abundance vary between morning and afternoon as an example of how NGS can be used to answer biological or ecological questions. If the statistical models used to analyse these data from the NGS method work well, then NGS could be an effective method to use in assessing pollen preferences of honeybees around AESs. To look at diurnal patterns in species presence and abundance, the samples were separated into morning and afternoon groups and compared. The presence/absence data were tested for significance between morning and afternoon using generalised linear models (family=binomial). The abundance data were log-transformed to convert the data from a Poisson distribution to a normal distribution and then tested using two-tailed paired t-tests. Communities of species composition in the morning and afternoon were compared using a Non Metric Multi-Dimensional Scaling (NMDS) analysis. The results of this ordination were tested for significance using an Adonis test.

2.2 Pollen preference methods

To assess the relative pollen preferences of honeybees for *P. tanacetifolia*, two field experiments were conducted along with one in a greenhouse.

2.2.1 Pollen collection when *Phacelia tanacetifolia* is present

The first field experiment sought to answer the research question of to what extent do honeybees forage on *P. tanacetifolia* for pollen in an agricultural landscape.

Setting

P. tanacetifolia was sown in strips 8 meters long and 1.5 meters wide every two weeks to ensure a constant blooming period of at least 6 weeks for pollen collection. The honeybee hives used for this

experiment were located about 25 meters from the strips of flowers. Two healthy hives with queens of similar age (determined by the beekeeper) were chosen, as the age of the queens affects the pollen demand of the colony and thus pollen foraging.

Pollen Collection

Pollen traps were installed on both honeybee hives and were again used to collect pollen. Pollen was collected every day that it was not raining and the temperature was at or above 14°C (a total of twelve days), as honeybees do not leave the hive to forage when it is cold or raining (Dimou et al. 2006). The main entrances of the pollen traps were closed at 11:00 and opened again at 13:00 to collect all of the pollen the foraging bees brought back to the hive during this 2-hour period. This time period was selected because honeybee foraging activity is high during the middle of the day (García-García et al. 2004), and this period was likely to include both those species that were foraged more heavily in the morning and those species that were foraged more heavily during the afternoon. Temperature, wind speed, wind direction and cloud cover were recorded halfway through the pollen collection at 12:00 as these conditions have been shown to have an effect on pollen foraging by honeybees.

All the pollen pellets from each time period were collected in 25 mL containers and stored in -20°C freezers to ensure that no fungi would grow on the pollen samples.

Species identification from pollen samples

The pollen pellets in each sample were weighed and counted. The purple pellets were separated from the others and placed in individual tubes, as they could potentially be *P. tanacetifolia* pollen (*P. tanacetifolia* pellets are purple). DNA was extracted from each of these purple pollen pellets to determine whether or not they were *P. tanacetifolia*. The same GF1 DNA extraction kit (Vivantis, Life Technologies) that was used for the NGS pollen identification methods was used for DNA extraction.

DNA was extracted, stored, and amplified by PCR using the protocol and primers described above (Section 2.1).

Gel electrophoresis was used to check if DNA was present in the PCR products, each of which was run through a 1.5% agarose gel to confirm the success of the PCR reaction. The gel was checked for bands, indicating the presence of DNA. The positive and negative controls were checked as well; if there was a band present for the negative control, then this indicated contamination of the samples.

If no contamination was present, then the PCR products were sequenced using standard Sanger sequencing in the forward and reverse direction. The resulting sequences were queried against the

GenBank database to determine what species they were. The data were analysed for proportion of *P. tanacetifolia* pollen pellets in each sample.

2.2.2 Observations of honeybees foraging for *Phacelia tanacetifolia*

To complement the pollen collection experiment, honeybees foraging on *P. tanacetifolia* were recorded. In this way, use of *P. tanacetifolia* by honeybees was examined from both the hive perspective through the pollen brought back to the hive from foragers and from the field perspective through the foragers' use of it in the field.

The strips of *P. tanacetifolia* from pollen collection experiment were used to make observations of honeybees on the flowers at 10:00, 12:00, and 15:00. At 9:00, there was often dew on the *P. tanacetifolia* flowers, so the honeybees would not forage on them. Therefore, 10:00 was chosen as the starting time for morning observations. The strips of P. *tanacetifolia* were monitored for honeybees for 5 minutes at the three times. The number of honeybees foraging for nectar and the number foraging for pollen were recorded. Twenty flowers of *P. tanacetifolia* were chosen at random and the amount of pollen on their anthers as well as their relative age were recorded. The amounts of pollen on anthers of the *P. tanacetifolia* flowers were scored based on the amount of visible pollen covering the anthers; the scoring scale is shown in Table 2.2.1. The ages of the *P. tanacetifolia* flowers were scored using the four stages of *P. tanacetifolia* flower longevity defined by Williams (1997); the scoring scale is shown in Table 2.2.2

Table 2.2.1 Scores for amount of pollen on *P. tanacetifolia* flowers

Amount of Pollen	Score
No visible pollen (0%)	0
Small amount of visible pollen (25%)	1
Some visible pollen (50%)	2
Large amount of visible pollen (75-100%)	3

Table 2.2.2 Scores for age of *P. tanacetifolia* flowers

Maturity of Flowers	Score
Just-opened flower (Stage 1: curled filaments and style)	1
Mid-age flower (Stage 2: filaments uncurled and petals at about 60°)	2
Mid-age-old Flower (Stage 3: petals at about 20-60°, styles longer than filaments)	3
Older Flower (Stage 4: petals closing, some anthers may have fallen off filaments)	4

Weather conditions of temperature, wind speed and direction, and cloud cover were recorded at each observation time.

Since the areas where the *P. tanacetifolia* was blooming changed throughout the experiment as some areas started to go to fruit and others started flowering, honeybee counts were divided by the area of *P. tanacetifolia* monitored during that time period. A mixed effects linear model was used to test for significance between the number of honeybees foraging for nectar and the number foraging for pollen at the different times of day. A mixed effects linear model was also used to test for the relationship between honeybees foraging for pollen and the amount of pollen on the flowers at different times of day and again to test for the relationship between overall number of foraging honeybees and maturity of flowers at different times of day.

2.2.3 Nectar volumes in *Phacelia tanacetifolia* flowers

The nectar volumes of *P. tanacetifolia* were recorded to see if the nectar volumes, and thus how attractive the *P. tanacetifolia* flowers can be to honeybees, fluctuate at different times of day. This experiment was done to answer the research question of what are the diurnal patterns of nectar flow in *P. tanacetifolia*. Trays of *P. tanacetifolia* were grown in a heated greenhouse with a 14 hr photo period. Once the plants were blooming, nectar volumes were measured using micro capillary tubes of fixed diameter to suck the nectar from the bottom of the flower. Then the amount of nectar in the tubes was quantified by measuring the length of liquid in the tubes, with each millimetre being equivalent to 1 microlitre. This was done for ten flowers at 10:00, 12:00, and 15:00 using micro capillary tubes and digital callipers to measure the length of the nectar in the tubes. At each time, the temperature in the greenhouse was also recorded. These three times were chosen to match the observational data collected at the same times on honeybees foraging on *P. tanacetifolia*. This process was repeated for ten days. A mixed effects linear model was used to test for differences in nectar volumes between the three times of day.

Chapter 3

Results

3.1 Pollen identification results

To answer the first and second research questions of how do pollen identification techniques compare against each other and to the proposed alternative of Next Generation Sequencing (NGS), Table 3.1.1 was created.

Table 3.1.1 Comparison of pollen identification methods based on time, cost, and accuracy

	Time	Cost	Accuracy	References
Visual	Low	Low	Low	Barth et al. 2009; Stanley & Linskens 1974; Arita et al. 1989
Microscopy	High	Moderate	Moderate	Barth et al 2010; Erdtman 1943
Flow Cytometry	Moderate	High	Unknown	Kron et al. 2014; Kron & Husband 2012)
Infrared Spectroscopy	Moderate	High	Unknown	Gonzalez-Martin et al. 2007; Gottardini et al. 2007
Computer Image recognition	Unknown	Unknown	Unknown	Marcos et al. 2014; Mitsumoto et al. 2009
DNA Sanger sequencing	High	High	High	Galimberti et al. 2014; Bruni et al. 2015
NGS	High	High	Moderate	

Time and cost values were based on personal observation and what was found in the literature. Comparing the accuracy of each of these methods experimentally was beyond the scope of this study. However, from the literature, the accuracy of some of the methods was deduced. Due to the novelty and limited use of some pollen identification methods to-date, some of the values were unknown. The visual categorisation method was thought to have the lowest time requirements, cost, and accuracy. Microscopy was thought to have the next lowest cost, and its time requirements and accuracy were high and moderate, respectively. DNA Sanger sequencing was thought to have the highest accuracy, as DNA barcoding would be able to identify individual pollen samples to species. Some of these methods would have low time requirements and costs if the number of samples to be processed was small (i.e., Microscopy and DNA Sanger Sequencing). For the purpose of creating Table 3.1.1 however, the values were estimated for a sample size similar to the one used to assess the NGS method (sample size =37).

Table 3.1.2 Comparison of pollen identification methods using additional criteria

	Use of Specialised Equipment	Use of Potentially Harmful Agents	Calibration Required	Reference Library Required	Level of Training or Experience	Post Processing Required	Ability to Quantify Uncertainty	References
Visual	No	No	No	Yes	Low	None	No	Barth et al. 2009; Stanley & Linskens 1974; Arita et al. 1989
Microscopy	No	Yes (strong acid to treat pollen grains)	No	Recommended	High	None	No	Barth et al. 2010; Erdtman 1943
Flow Cytometry	Yes	No	Yes	Yes	Mid	None	Yes	Kron et al. 2014; Kron & Husband 2012
Infrared Spectroscopy	Yes	No	Yes	Yes	High	None	Yes	Gonzalez-Martin et al. 2007; Gottardini et al. 2007
Computer Image recognition	Yes	No	Yes	Yes	High	None	Yes	Marcos et al. 2014; Mitsumoto et al. 2009
DNA Sanger sequencing	Yes	Yes (liquid nitrogen)	No	Recommended	Mid	None	Yes	Galimberti et al. 2014; Bruni et al. 2015
NGS	Yes	Yes (liquid nitrogen)	No	Recommended	Mid	High	Yes	

The values in Table 3.1.2 were based on personal observation and the literature.

3.1.1 Results of the NGS method

For the third research question on how well NGS can detect species presence and abundance, the triplicates (subsamples) within each sample group were compared against each other. First the subsamples within the same sample group were compared for species presence. Species detected in both of the compared subsamples was interpreted as a positive detection, and species detected in only one of the subsamples compared was a missed detection. Within each subsample comparison, the percent detection was calculated and then averaged for each sample group. Table 3.1.3 shows the subsample comparisons for species detection in the first three sample groups.

Table 3.1.3. Detection percentage of species presence in subsamples from samples 1-3

Sample	Subsample comparison	Detection Percentage of Present Species	Average Species Detection per Sample
1	1a 1b	67%	
1	1b 1c	50%	72%
1	1a 1c	100%	
2	2a 2c	33%	
2	2b 2c	75%	61%
2	2a 2c	75%	
3	3a 3b	25%	
3	3b 3c	80%	43%
3	3a 3c	25%	

Generalized linear models were used to assess the consistency of species abundance between subsamples within the same sample group. The slope values from the summary statistics were used to compare how well the species abundances between two triplicate subsamples correlated to each other. A slope value closer to zero represented a small or negligible difference between species abundances between subsamples. Figure 3.1.1 shows the mean slope values for each sample group calculated from the generalised linear models.

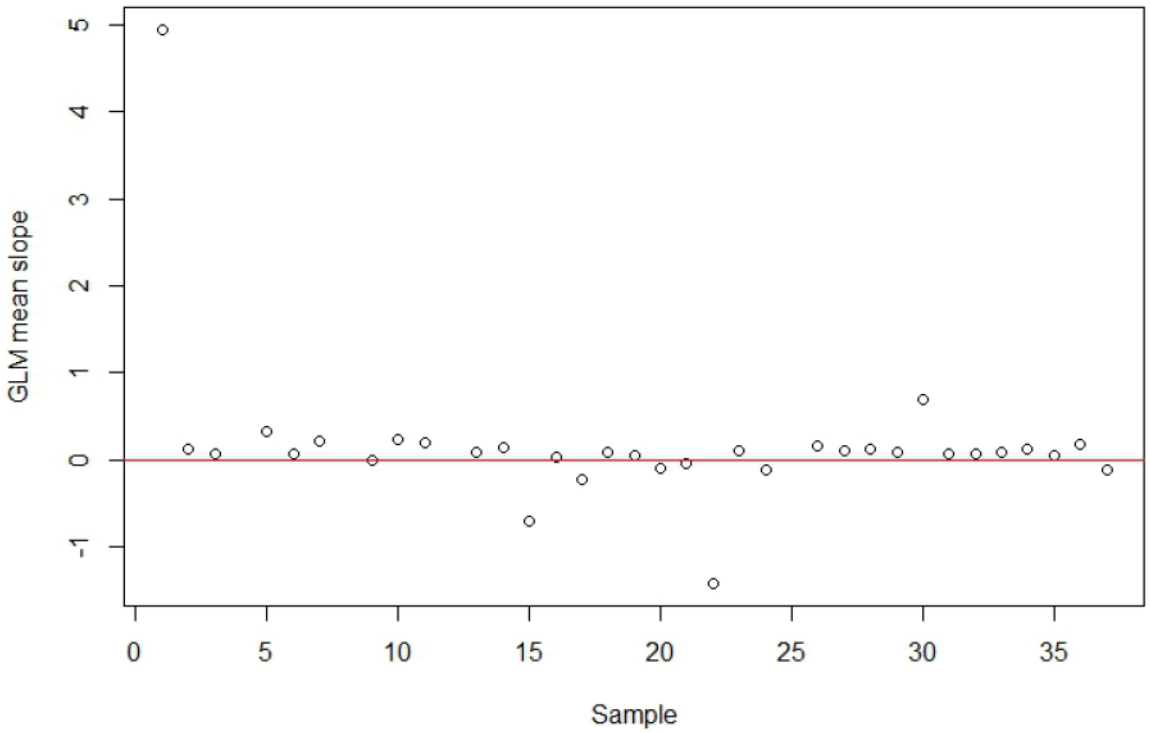

Figure 3.1.1 Mean slope values calculated from generalised linear models for each sample group. A value closer to zero represents a higher amount of similarity between species abundances in subsamples.

While the mean slope values are relatively close to zero, the p values from the generalised linear models revealed that 72 out of the 105 comparisons between subsamples were significantly different from each other.

The data was also analysed to test whether a higher species abundance in subsamples was related to a higher detection percentage in sample groups using a linear model. The results of the linear model showed that there was a significant relationship between species detection and species abundance when the percentage of species detection was 67% and 100% (p value= 2×10^{-16} for both the 67% and 100% detection percentages; adjusted $R^2 = 0.5425$).

Species Presence and Abundance in the Morning and Afternoon

Species presence and species abundance data were also used to test using the NGS method to answer the simple biological question of does species presence and species abundance change differ between the morning and afternoon samples. As the dataset contained days in which many of the species were neither present in the morning or afternoon (double zeros), only the species commonly found in the dataset and present in either the morning or afternoon were used to compare between

morning and afternoon. Species present in at least 30% of the samples (morning or afternoon) were considered common and used for this analysis. Species presence data were analysed using generalized linear models. Abundance data were analysed using tow-tailed paired t-tests. These results are shown in Table 3.1.4.

Table 3.1.4 Species presence and species abundance p values from GLMs and paired t-tests

Species	Presence p value	Abundance p value
Raphanus sativus	0.3016	0.2825
Trifolium pratense	1.00	0.8263
Trifolium repens	0.423	0.1203
Brassica spp.	0.654	0.1995
Taraxacum mongolicum	0.0563	**0.0030**
Trifolium squamosum	0.2604	0.2353

Only the species *Taraxacum mongolicum* (common name Dandelion) was found to have a significant relationship between abundance in the morning and abundance in the afternoon. As shown in the Figure 3.1.2 below, *T. mongolicum* significantly decreased in abundance between in the morning and afternoon.

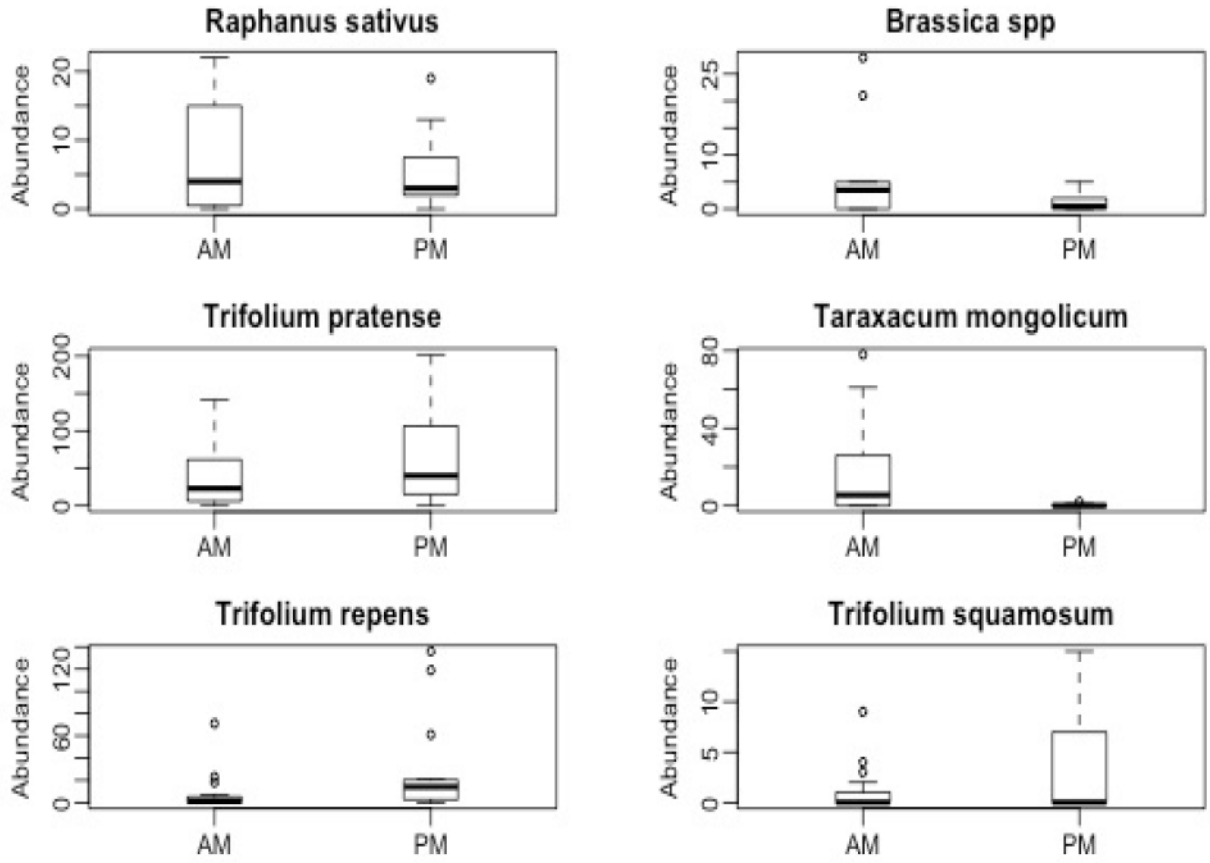

Figure 3.1.2 Species Abundance (represented by number of sequences in the combined samples) in the morning and afternoon. Only *Taraxacum mongolicum* had a significant relationship between morning and afternoon abundance.

Additionally, species composition between the morning and afternoon was compared using NMDS (Nonmetric Multi-dimensional scaling) ordination. For this analysis, all species, even the very rare ones, were included. The Adonis test was not significant (Adonis method "Bray", stress=0.196, p value = 0.092, R^2= 0.0478, 0.9522 for AM PM and Residuals respectively), meaning that there was not a significant difference in species composition between the morning and afternoon. Figure 3.1.3 below shows the NMDS ordination plotted on two axes as a way to visualise species communities in the morning and afternoon samples. There are no significant patterns in the data community composition.

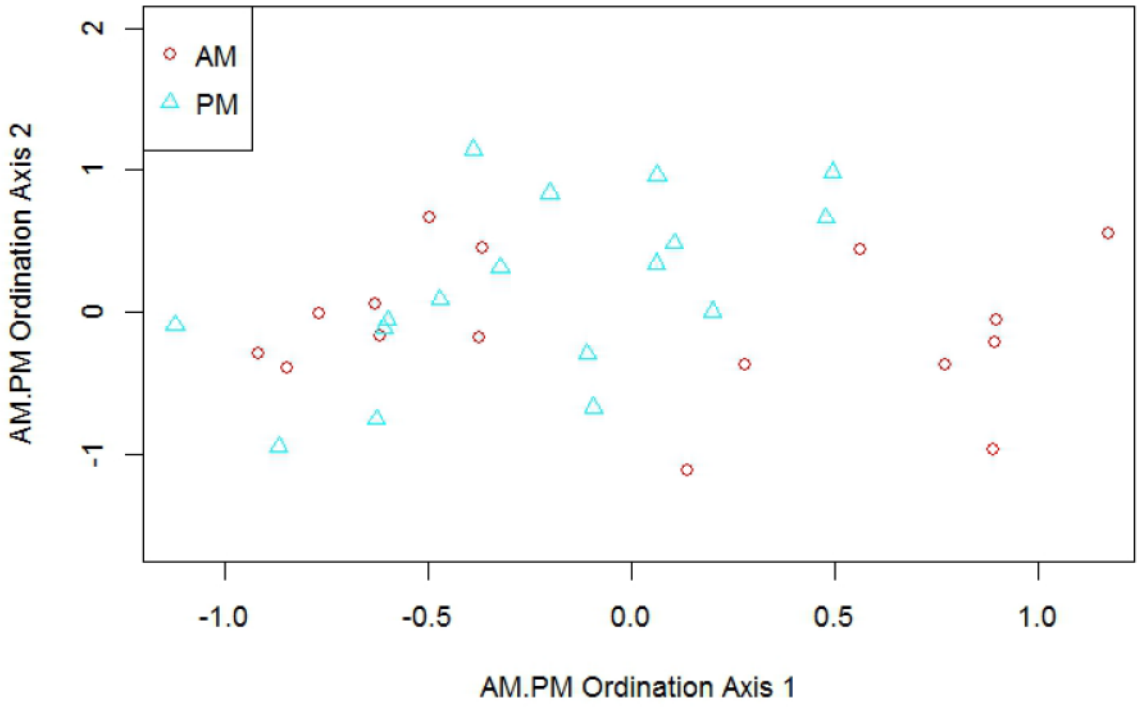

Figure 3.1.3 Ordination plot of the species composition between the morning and afternoon.

3.2 Pollen preference results

3.2.1 Pollen collection from honeybee hives

Only one *P. tanacetifolia* pollen pellet was found in the pollen samples, or 1 out of 23, 431 total pollen pellets collected. No statistical tests for significance were run on these data as the number of pollen pellets was very negligible.

3.2.2 Observations of honeybees on *Phacelia tanacetifolia* in the field

First the data were compared to test if there was a difference between honeybees foraging for pollen and honeybees foraging for nectar using a two-tailed paired t-test. With a p-value = 6.73×10^{-16}, the number of honeybees foraging for pollen was significantly less than the number of those foraging for nectar. Figure 3.2.1 illustrates the number of honeybees foraging for pollen and those foraging for nectar at different times of day.

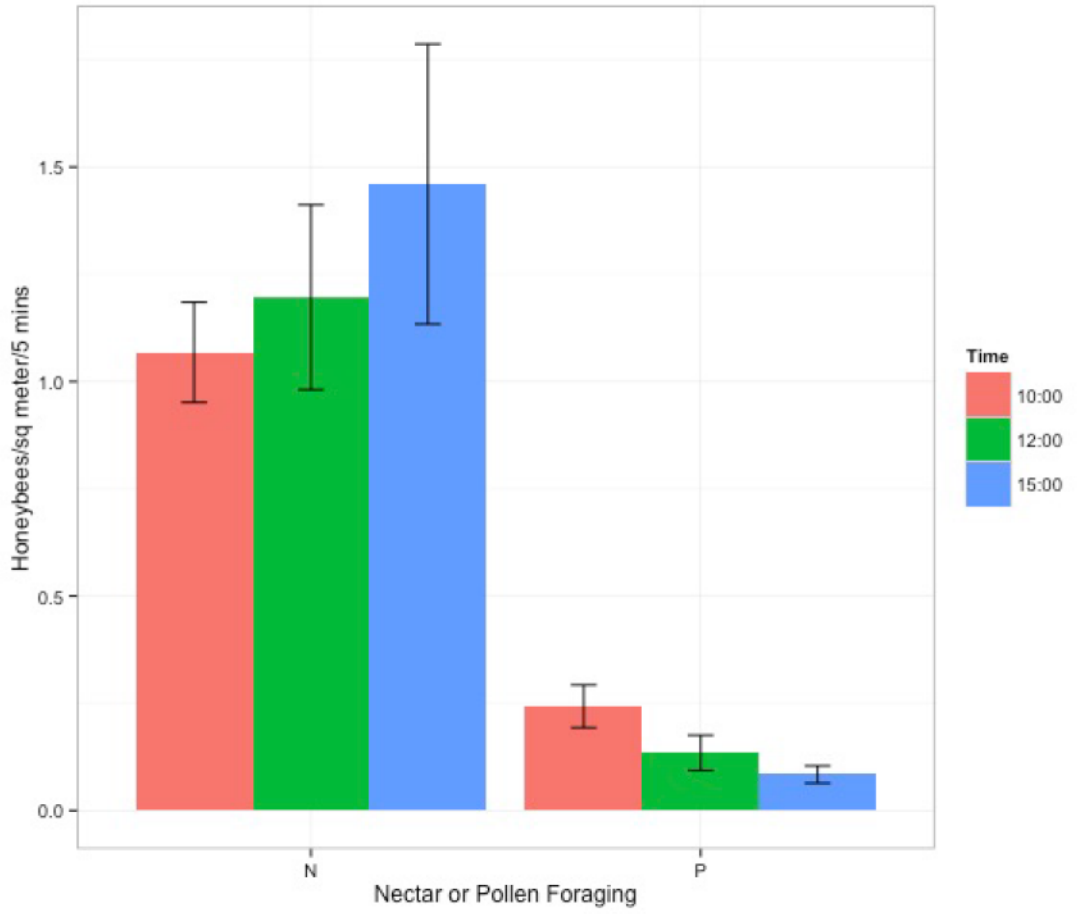

Figure 3.2.1. Number of honeybees foraging for nectar (N) and foraging for pollen (P) at 10:00, 12:00, and 15:00. Error bars represent respective standard errors.

The boxplots in Figure 3.2.2 below were used to visualise the relationships between foraging honeybees, pollen amount, and flower maturity. The boxplot on the left of pollen foraging honeybees and pollen amount shows that there was a trend towards decrease in number of honeybees foraging for pollen when there was 50% pollen on the *P. tanacetifolia* anthers (Score = 2). The boxplot on the right shows that the median number of total honeybees, both pollen and nectar foraging, did not vary between the four stages of flower maturity.

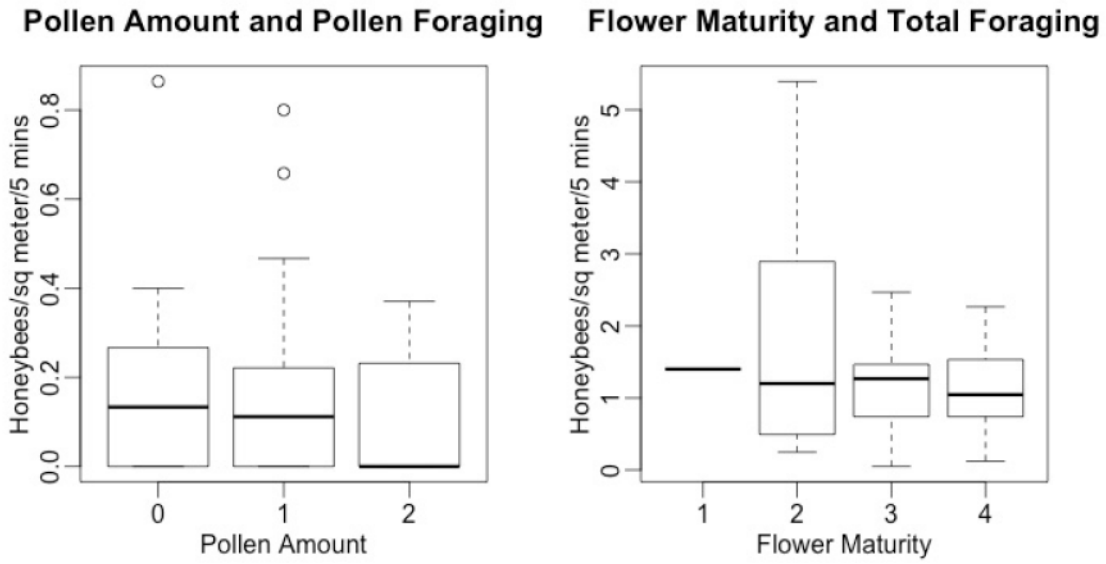

Figure 3.2.2 Boxplots showing the relationship between number of honeybees foraging for pollen and the amount of pollen found on the *P. tanacetifolia* anthers (left) and showing the relationship between total honeybees foraging and flower maturity stage (right).

Linear mixed effects models were used to test if differences in honeybee foraging numbers were related to time of day, pollen amount on the *P. tanacetifolia* flowers, or flower maturity. These models accounted for the random effects in the data due to day on which the data were taken. The linear mixed effects models were bootstrapped for the data on pollen foraging honeybees, nectar foraging honeybees and total foraging honeybees to determine the 95% confidence intervals around each of the coefficient values generated from the model. Figures 3.2.3, 3.2.4, and 3.2.5 below show the coefficient values and corresponding confidence intervals in relation to the intercept of the models. Those coefficient values which had confidence intervals which overlapped zero were taken to indicate that there was no significant effect of that variable on the foraging honeybees. This is because a possible zero-value coefficient in the linear model indicates that the corresponding variable for the coefficient can be regarded as having no significant effect on the data.

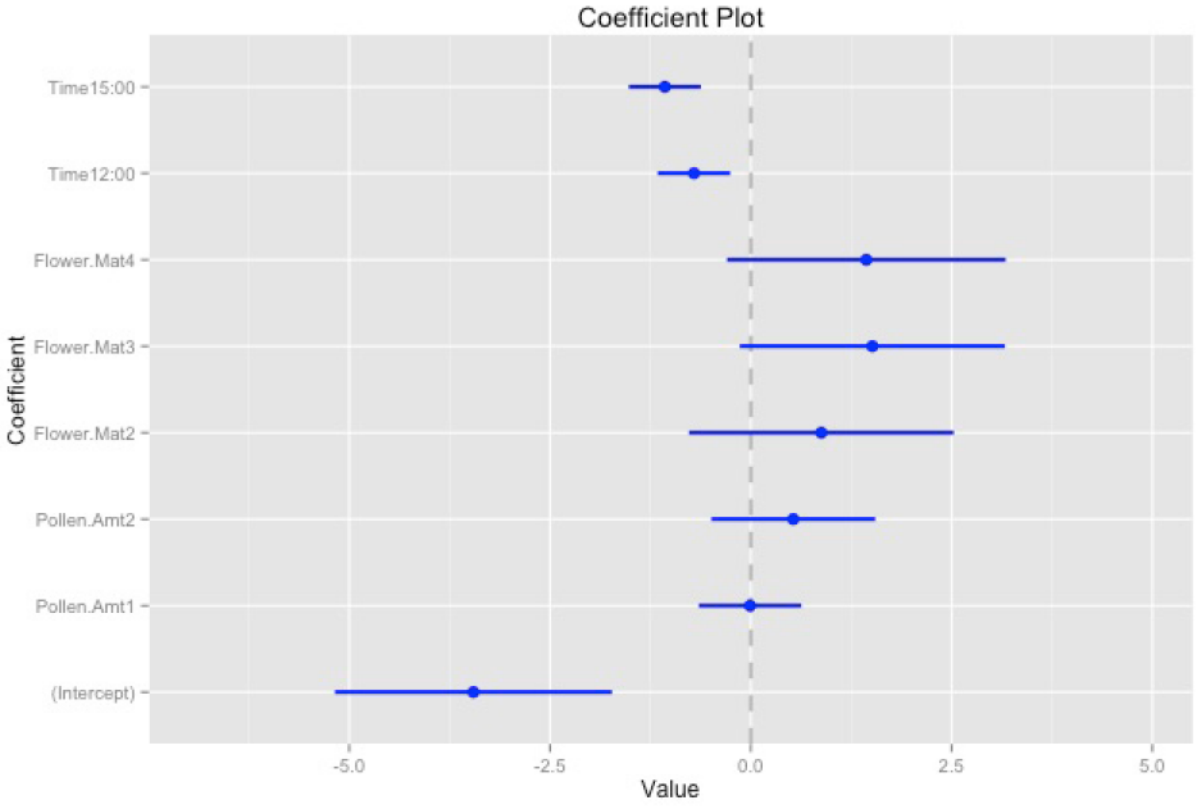

Figure 3.2.3 Coefficient Plot with 95% confidence Intervals for pollen foraging honeybees. The intercept refers to the baseline values used for comparisons (Pollen Amount =0, Flower Maturity=1, and Time =10:00).

In Figure 3.2.3, coefficient values whose 95% confidence intervals did not overlap with zero indicate a significant effect on the number of pollen foraging honeybees. In relation to 10:00, the number of pollen foraging honeybees decreased at 12:00 and 15:00. Results from a Likelihood Ratio Test however, show that time did not have a significant relationship with the number of honeybees foraging for pollen (test=ANOVA, df=4,6, χ^2= 2.132, p value=0.344). Neither pollen amount nor flower maturity had a significant effect on the number of pollen foragers (results from a Likelihood Ratio Test test=ANOVA, df=6,10, χ^2 = 5.648, p value=0.227; test=ANOVA, df=5,10, χ^2 = 7.749, p value=0.171 for Flower Maturity and Pollen Amount respectively). Although the above Figure 3.2.2 showed that there was a trend in decreasing pollen foraging honeybees at 50% pollen amount, since the confidence interval overlaps with zero, the trend was not significant.

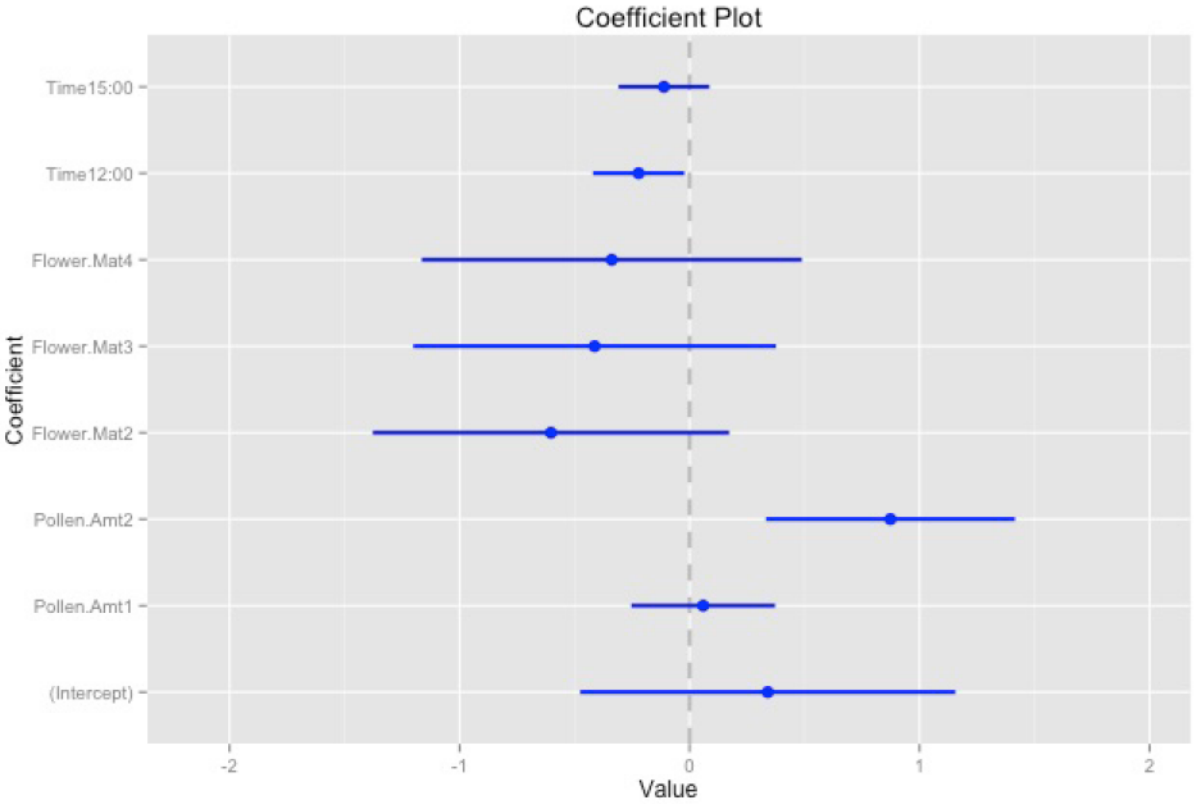

Figure 3.2.4 Coefficient plot with 95% confidence intervals for nectar foraging honeybees. The intercept refers to the baseline values used for comparisons (Pollen Amount =0, Flower Maturity=1, and Time =10:00).

In Figure 3.2.4, coefficient values whose 95% confidence intervals did not overlap with zero indicate a significant effect on the number of nectar foraging honeybees. In relation to 10:00, the number of pollen foraging honeybees decreased a significant amount at 12:00, however there was not a significant relationship between nectar foragers at 10:00 and at 15:00 as confirmed by a Likelihood Ratio test (test=ANOVA, df=5,10, χ^2 = 4.984, p value=0.418). In relation to no visible pollen on the *P. tanacetifolia* anthers, there was a increase in nectar foraging honeybees when there was some visible pollen on the anthers (Score=2), although this relationship between pollen amount and honeybees foraging for nectar was not significant (test=ANOVA, df=5,10, χ^2 = 1.913, p value=0.861). There was no significant relationship found for the flower maturity variables, as their coefficient confidence intervals overlapped with zero, indicating a possible coefficient value of zero. A Likelihood Ratio test also confirmed a nonsignificant relationship (test=ANOVA, df=6,10, χ^2 = 4.988, p value=0.289).

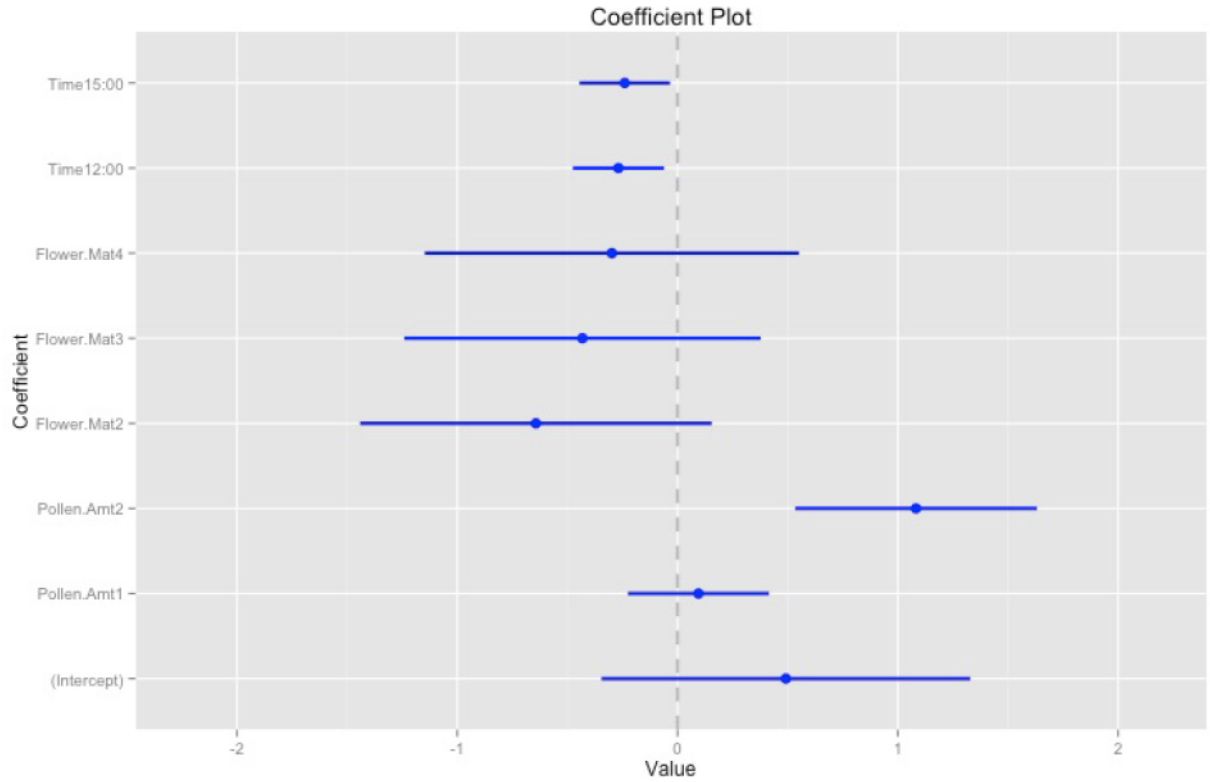

Figure 3.2.5 Coefficient plot for total foraging honeybees. The intercept refers to the baseline values used for comparisons (Pollen Amount =0, Flower Maturity=1, and Time =10:00).

In Figure 3.2.5, coefficient values whose 95% confidence intervals did not overlap with zero indicate a significant effect on the number of nectar foraging honeybees. In relation to 10:00, the total number of foraging honeybees decreases at 12:00 and 15:00, although this amount is not significant (test=ANOVA, df=5,10, χ^2 = 7.141, p value=0.210) . In relation to no visible pollen on the *P. tanacetifolia* anthers, some visible pollen increases the number of total honeybees foraging on *P. tanacetifolia*. However, there is no significant relationship between pollen amount and the total number of foraging honeybees (test=ANOVA, df=5,10, χ^2 = 2.862, p value=0.721). There was not a significant relationship between flower maturity and total number of foraging honeybees (test=ANOVA, df=6,10, χ^2 = 7.004, p value=0.136).

3.2.3 Nectar volumes of *P. tanacetifolia*

Figure 3.2.6 displays the relationships between nectar volumes and time of day.

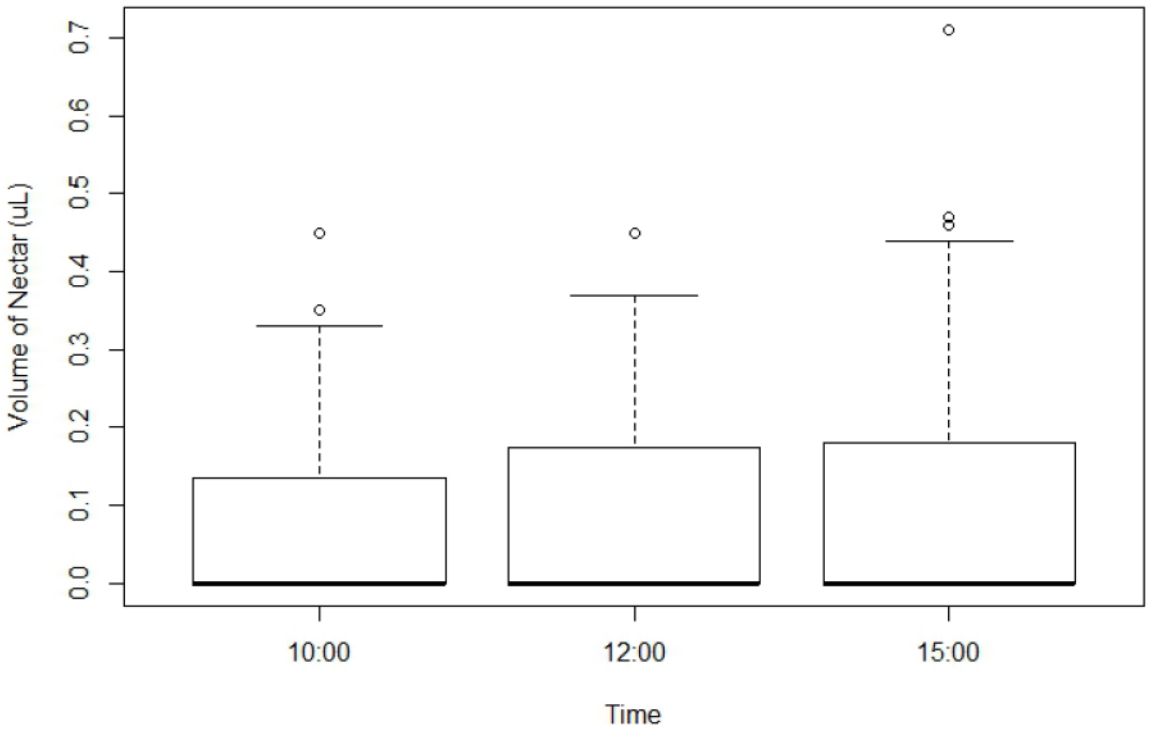

Figure 3.2.6 The relationship between volume of nectar and time of day. The medium volumes, shown by the bold black line in the boxplots, do not vary between 10:00, 12:00, and 15:00.

Linear mixed effects models were used to test if time of day and temperature were related to nectar volumes in *P. tanacetifolia*. Using the linear mixed effects models, a Likelihood Ratio Test was used to determine if the isolated effect of time of day was related to nectar volume. There was found to be no significant relationship (test=ANOVA, df=5,6, χ^2 = 1.484, p-value = 0.223). Another Likelihood Ratio Test was used to determine if the isolated effect of temperature was related to nectar volume. It also was found to not have a significant relationship (test=ANOVA, df=4,6, χ^2 = 4.746, p-value = 0.0932).

Chapter 4

Discussion

4.1 Pollen identification

4.1.1 Methods of pollen identification

Based on the comparison tables and a literature review (Tables 3.1.1 and 3.1.2) of the pollen identification methods, no one method was the best overall option. From this synthesis of the literature, it can be concluded that the efficiency and feasibility of these methods heavily depend on the project's aim, sample size, budget, and available resources. The visual categorisation method was the most user-friendly as it does not require any specialised equipment or extensive training. While it is a very low-cost option, it also can be very inaccurate. The conventional method, microscopy, offers the advantage of being moderately accurate and not requiring any specialised equipment. However, its biggest disadvantage is that it requires a high level of training or experience to be able to identify correctly the pollen grains to plant genus or species level. If someone with experience in pollen identification using microscopy is involved in the research project, then this method may be the most cost and time effective.

The flow cytometry, DNA Sanger sequencing, infrared spectroscopy, and computer image recognition methods all require the specialised equipment of a cytometer, sequencing machine, spectrometer, and computer software, respectively. Flow cytometry, infrared spectroscopy, and computer image recognition require calibration, meaning that the machines would have to be trained to sort outputs into known species groups. These three methods would also require a reference library to calibrate the machines, so that the machines could recognise all of the possible species that could be found in a pollen sample as being unique. These processes would add on a large amount of additional time. As these methods are still being developed, the methods would not be suitable for projects with an unknown mix of species or a large number of species. Therefore, these methods would not be recommended for use in assessments of Agri-Environment Schemes (AESs) looking at on which plant species honeybees forage for pollen.

DNA Sanger sequencing is a more accurate method and would be well suited for small sample sizes since each pollen pellet would need to be processed individually to ensure correct identification of species. With large sample sizes however, DNA Sanger sequencing would be very time-consuming and expensive. As another type of DNA sequencing, NGS (Next Generation Sequencing) was investigated in this project with the idea that it could have the potential for being accurate as well as

being able to process large samples of pollen. This next section will discuss results and their implications for using NGS as a pollen identification method.

4.1.2 NGS as an alternative pollen identification method

The NGS results were first analysed for the ability of NGS to determine species presence and species abundance. The results of the analysis examining the accuracy and consistency of the NGS method within sample groups showed that the NGS method is not the most accurate for detecting species, with a mean presence detection percentage of species between subsamples about 56%. The results of the generalised linear models testing if there was a difference in abundance detections between subsamples showed that most of the subsample comparisons had significant differences, with 72 of the comparisons between subsamples being significantly different from each other out of the 105 comparisons. This means that within the same sample group, most of the subsamples had significantly different species' abundances from each other. The analysis also used the slope coefficient outputs from the generalised linear models to assess scale of the differences between subsamples, and most of the slope differences were not very large and different from zero (a slope value of zero meaning that there was no difference between the subsamples compared). This indicates that although most of the subsamples within the sample groups were comprised of significantly different species abundances, these differences were not very large. Thus, with larger subsample sizes (five or more), abundance data between subsamples may be more similar and statistically equal, meaning that abundance data inferred from NGS pollen identification methods is accurate enough for analysis.

Looking at the analysis of detection percentages of individual species, there was a higher rate of detection of species in larger abundances. This means that NGS as a method may be suitable for detection of species when it is known that honeybees will forage on them in high abundance for pollen. Due to the capability of NGS to process very large sample sizes, it may be a more cost- and time- effective method than other methods for surveying the species on which honeybees forage for pollen very often. For small sample sizes requiring a high level of accuracy and consistency between samples, NGS would not a suitable method to use and DNA Sanger sequencing may be more suitable.

One explanation for these results is that was a high number of sequencing errors with homopolyers in the primer sequences, or sequences with more than one of the same basepair in a row. Deletion and insertion errors in the homopolymers of the primer sequences resulted in a number of sequences being excluded from the data analysis as they were not able to be sorted into sample groups. Errors in homopolymer sequencing is not uncommon for 454 NGS sequencing methods (Shendure & Hanlee 2008). However, in a study using 454 to sequence microbial samples with known amplicon composition, Quince et al. (2009) found that 454 was still a reliable method for determining

diversity and abundance if the sequences were de-noised and chimera sequences were removed. This project did not follow the exact pipeline recommended by Quince et al. (2009), but as the resulting sequences were de-noised and chimera sequences were removed, the conclusions from that study can be applied here. While this process removed a large number of sequences in the data, the remaining sequences can be assumed to be accurate.

The results of this study on assessing how effective NGS would be for identification of honeybee-collected pollen showed that this method is not very accurate for detecting species with low abundance in the samples and nor is it accurate at detecting exact species abundance represented in the samples. These results indicate that the method still has some more refinement required to increase its accuracy and consistency, although it remains a promising option for pollen identification in the future, especially for large sample sizes.

Species presence and abundance in the morning and afternoon

The sequence data for subsamples were combined into their respective sample groups and used to answer the question of whether the presence and abundance of the species on which honeybees forage changes between the morning (10:00-11:00) and afternoon (13:00-14:00). The results of this analysis found that only *Taraxacum mongolicum* (common name Dandelion) was represented in higher abundance in the morning than in the afternoon, meaning that honeybees forage on *T. mongolicum* more in the morning compared to the afternoon. None of the other species were present or abundant in significantly different amounts between the morning and afternoon. Only the species commonly represented in the samples were compared for this analysis because the other 23 species were found only in a few samples and in small numbers. This may be because NGS as a method has a higher likelihood of detecting abundant species over the rare ones in samples.

The results of the NMDS ordination plot showed that species composition, which included all of the species in the data analysis, was not significantly different between the morning and afternoon (p value = 0.092). These results support the species presence and abundance analyses.

Similar studies in the literature support these results on the change, or lack thereof, of species presence and abundance between the morning and afternoon in honeybee-collected pollen samples. García-García et al. (2004) studied the variation in collected pollen weights for the early morning, morning, early afternoon, afternoon, and late afternoon. The study found that while some species were more foraged on by honeybees at certain times of day, many species were foraged on by honeybees at similar rates throughout the day, similar to the results of this present study. Knowing when honeybees tend to forage on particular species for pollen can be helpful for AES assessments. If certain plant species are foraged on more in the morning, then this information is important to know when designing the experiment to assess AESs.

Limitations and future work

Along with the sequencing errors, another limitation of this study was that there was not an established reference library, or a list of sequences from all possible species that could be found in the samples. The sequences were blasted against the GenBank database, and while this database would likely have the vast majority of the species present in the samples, there is a small chance that a few of the native New Zealand species which could have been in the samples were not on GenBank and thus the sequences would have been incorrectly matched or discarded. It is recommended for future studies to use a reference library to ensure that all sequences would be correctly matched to species.

For future research, it is recommended to examine Illumina, a newer type of NGS, as another method for pollen identification. The Illumina method may have less errors when sequencing pollen, which would result in more sequences in the data to be analysed. Richardson et al. (2015) used Illumina for a metabarcoding study on honeybee-collected pollen and compared the species found to be present in the results from the Illumina sequencing to species identified using microscopy. This study found that the Illumina NGS method was able to identify more of the species that were less abundant in the samples than microscopy was able to, suggesting that Illumina may be a more suitable NGS method for pollen identification than 454 NGS. However, Richardson et al. (2015) did not assess the consistency of the Illumina method as this current study has done; a comparison assessing the consistency and accuracy of both of these NGS methods should be done. Additionally, the use of other primers to amplify pollen DNA sequences should be explored and compared to ITS. Galimberti et al. (2014) used the primers rbcl and trnH-psbA and had high amplication success. Using these primers in conjunction with ITS may increase NGS's ability to detect species presence.

4.2 Pollen preference

4.2.1 Honeybees' Use of *Phacelia tanacetifolia*

Only one *P. tanacetifolia* pollen pellet was found in all of the collected pollen pellets, which indicates that honeybees' use and therefore pollen preference of *P. tanacetifolia* is negligible. The results of the pollen collection experiment are complemented by the experiment observing honeybees' foraging behaviour on the *P. tanacetifolia*. These results found that honeybees forage for nectar significantly more than they forage for pollen at the times of 10:00, 12:00, and 15:00. Combining these results with the results of the pollen collection experiment shows that while honeybees may not prefer the pollen of *P. tanacetifolia*, they still use the flowers for nectar. These results are supported by a study done by Williams and Christian (1991) which found that most visits by honeybees on *P. tanacetifolia* were for nectar rather than pollen (78% honeybees foraging for nectar and 22% foraging for pollen).

The results of the foraging behaviour observations also showed that pollen amount of the flowers' anthers had a significant effect on both nectar foraging and total honeybee foraging on *P. tanacetifolia*. While increased pollen amount did not have a significant effect on pollen foraging, there was a trend in increased pollen foraging and may not have been significant due to low numbers of pollen foragers in general. These results indicate that increased pollen amount of *P. tanacetifolia* flowers may make it a more attractive floral resource. While the pollen amount on individual flowers cannot be increased, a higher number of *P. tanacetifolia* plants would supply a higher pollen amount in general, which may make the *P. tanacetifolia* as a floral resource more attractive to honeybee colonies.

From the observational results, it was found that pollen foraging and total foraging on *P. tanacetifolia* was highest at 10:00 and then decreases at 12:00 and 15:00. One explanation for pollen foraging decreasing throughout the day is that pollen is depleted in the morning and then is in a much limited supply at mid-day and in the afternoon. However, *P. tanacetifolia* flowers will open throughout the day, especially in warmer weather (R. Sprague, observation), contradicting this explanation. Additionally, from the results of the greenhouse experiment on nectar volumes of *P. tanacetifolia* at 10:00, 12:00, and 15:00, it was found that nectar volumes did not change significantly between the three times. These results complement the study done by Williams (1997) that showed that nectar secretion of *P. tanacetifolia* is related to age of the flowers, with 4 hr-old and 7 hr-old flowers having the highest nectar secretion rates. Since the flowers of *P. tanacetifolia* open throughout the day, then they will secrete nectar throughout the day. Thus honeybees could continue to forage on the flowers in the morning and afternoon and not completely deplete their nectar resources. *Phacelia tanacetifolia*'s continued nectar production throughout the day provides another reason for its use as a floral resource in AES as long as there are enough plants to meet the demands of foraging honeybees (see Al-Ghamdi et al. 2014).

Another explanation for decreasing honeybee foraging activity throughout the day (and low foraging on *P. tanacetifolia* in general with a maximum foraging of only 5.4 honeybees per square meter) is that *P. tanacetifolia* is a limited resource in the environment. Because honeybees as a colony tend to prefer to forage on high-density, high-quantity flowers (Tautz 2008), the *P. tanacetifolia* may not have been a large enough resource to be very attractive to the honeybee colonies. Thus, while some individual bees would forage on *P. tanacetifolia*, these bees may have been recruited throughout the day to forage in more preferable areas. To test whether *P. tanacetifolia* is a limited resource or is of limited desirability, *P. tanacetifolia* should be planted as a mass flowering crop, as recommended by Westphal et al. (2003) as a way to benefit pollinators, to provide enough of it as a resource to be a viable attractant to honeybees. Then these same experiments of pollen collection and observing foraging behaviour on the *P. tanacetifolia* should be conducted to test for pollen preference and the

use of *P. tanacetifolia*. A recent study by Henry et al. (2012) planted a total of 1.2 square kilometres of *P. tanacetifolia* strips in a 5 kilometer radius around honeybee hives from which they collected pollen. The study found as much as 32% of the honeybee-collected pollen was from the *P. tanacetifolia* strips, which suggests that when there is enough of *P. tanacetifolia* in a landscape, honeybees will forage on the *P. tanacetifolia* flowers quite heavily for pollen. Therefore, *P. tanacetifolia* should be planted as a mass flowering crop to test this idea that it was a limited resource in this present study. If the results of the mass flowering crop experiment find that pollen preferences and use of *P. tanacetifolia* increase significantly when *P. tanacetifolia* is an abundant floral resource, then its pollen is a limited but preferred resource and the plants should be planted in large areas and higher quantities when they are used in AES or similar pollinator enhancement schemes.

Limitations

Both of these field experiments were only conducted in one season in one agricultural location, thus these results are very environmentally context dependent. More field experiments should be conducted in different seasons and different agricultural landscapes to verify or refute the results of this study. Furthermore, throughout the field experiments there were 12-24 honeybee hives near the field of planted *P. tanacetifolia*. This number of hives may have caused increased competition between honeybee colonies for resources. Since pollen was only collected from two of these hives, those two may have been out-competed from other hives to use *P. tanacetifolia* or *P. tanacetifolia* may have been too limited a resource to provide much pollen for any of the hives. To address this limitation, *P. tanacetifolia* should be planted as a mass flowering crop as mentioned above.

These field experiments also did not quantify all of the available floral resources and the available pollen per plant within the study area, as this quantification would allow the study to compare the size and number of available *P. tanacetifolia* plants to the size and number of surrounding floral resources. To address this limitation, a botanical survey of the surrounding area could be done in future studies to discover what other floral resources and how large an area of them were available to the foraging honeybees. Knowing how resource-abundant or resource-depleted a landscape was would help determine if *P. tanacetifolia* was a limited resource or a less desirable resource.

4.2.2 Implications for Pollinator Enhancement Schemes

Studies on *P. tanacetifolia* have shown that its pollen has a high protein content and that its pollen and nectar can improve the fitness and longevity of insects (Hickman & Wratten 1996; Laubertie et al. 2012). However, if honeybees only use the *P. tanacetifolia* for its nectar and not its pollen, then the honeybees are not benefiting from the *P. tanacetifolia* as researchers may have thought they would. This has serious implications for *P. tanacetifolia*'s use in these schemes for improving

pollinator fitness. More studies will need to be conducted in order to assess honeybees' preference for *P. tanacetifolia* when it is available in large amounts, such as a mass flowering crop.

Previous studies have exposed the lack robust scientific assessments of AES (Kleijn et al. 2006; Whittingham 2007), but these AES and pollinator health plans proposed by the US Federal government also lack monitoring and recording programmes as part of their enhancement measures. Both the Environmental Stewardship Handbook for AES and the US Pollinator Health Strategy 2015 recommend that farmers and landowners adjust the landscape measures to make the measures more effective at improving pollinator health (Natural England 2013a & 2013b; Vilsack & McCarthy 2015). However, neither of these programmes suggest specific ways in which farmers and landowners can monitor changes in pollinator health or activity before and after they enact the landscape enhancement changes. If these pollinator health policies could include specific guidelines for recording and monitoring, then changes that farmers and landowners make to the landscape measures would be more objective and evidence-based. Furthermore, the records kept on pollinator health and activity over time as the landscape measures were implement could be used to assess their effectiveness and success.

The research aim and design of the three experiments done in this project should be used as exemplars for future assessment measures of AES and similar schemes such as the US Pollinator Health Strategy 2015. The design offers different methods for evaluating honeybees' use of a flower, both through pollen collection and observations on foraging behaviour, which could be replicated in a variety of landscapes. The assessment measures used in this study to evaluate honeybees' use of *P. tanacetifolia* could easily be incorporated into a monitoring programme part of a pollinator health improvement government policy.

4.3 Conclusion

While these experiments did have limitations, their results contribute to understanding the advantages and limitations of pollen identification methods and to assessing AES and pollinator enhancement schemes. In spite of the errors and lower detection rate of species than was anticipated, the results of the NGS method assessment indicate that the method can detect presence well for the abundant species within pollen samples. Since NGS using 454 sequencing is thus a useful tool for pollen identification of species presence in large sample sizes, this method could be used in the future to assess how honeybees use the floral resources added to landscapes as part of pollinator enhancement schemes. If the NGS method identified the presence of the scheme's floral resources in the honeybee-collected pollen samples, then this suggests that those species were abundant enough to be detected by NGS and therefore the honeybees were using the floral resources relatively frequently.

The experiments examining honeybees' use of *P. tanacetifolia* found that the honeybees did not use the flowers much as a pollen resource although they did use it as a nectar resource. The results of this study laid the foundations for future research aiming to evaluate the key components of AES and pollinator enhancement schemes by providing a framework for the experimental design. Increased understanding of how pollinators use the high-quality floral resources added to landscapes and whether pollinators prefer these resources will lead to better-designed pollinator enhancement schemes and thus healthier pollinators.

Appendix A

Fusion Primers used for 454 Next Generation Sequencing (NGS)

A.1 Table of complete list of fusion primers (forward and reverse) for each subsample

Subsample	Complete Fusion primer Forward	Complete Fusion primer Reverse
1a	CGTATCGCCTCCCTCGCGCCATCAGTCTAGC GACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCAGA CGTCTGGCTGCGTTCTTCATCGATGC
1b	CGTATCGCCTCCCTCGCGCCATCAGTCTAGC GACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCAGT ACTGCGGCTGCGTTCTTCATCGATGC
1c	CGTATCGCCTCCCTCGCGCCATCAGTCTAGC GACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGAC AGCGAGGCTGCGTTCTTCATCGATGC
2a	CGTATCGCCTCCCTCGCGCCATCAGTCGCAC TAGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCAGT ACTGCGGCTGCGTTCTTCATCGATGC
2b	CGTATCGCCTCCCTCGCGCCATCAGTCGCAC TAGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGAC AGCGAGGCTGCGTTCTTCATCGATGC
2c	CGTATCGCCTCCCTCGCGCCATCAGTCGCAC TAGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGCG TGCTAGGCTGCGTTCTTCATCGATGC
3a	CGTATCGCCTCCCTCGCGCCATCAGTACGCT GTCTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGAC AGCGAGGCTGCGTTCTTCATCGATGC
3b	CGTATCGCCTCCCTCGCGCCATCAGTACGCT GTCTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGCG TGCTAGGCTGCGTTCTTCATCGATGC
3c	CGTATCGCCTCCCTCGCGCCATCAGTACGCT GTCTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGAT CTGTCGGCTGCGTTCTTCATCGATGC
4a	CGTATCGCCTCCCTCGCGCCATCAGACGAGT GCGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGCG TGCTAGGCTGCGTTCTTCATCGATGC
4b	CGTATCGCCTCCCTCGCGCCATCAGACGAGT GCGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGAT CTGTCGGCTGCGTTCTTCATCGATGC
4c	CGTATCGCCTCCCTCGCGCCATCAGACGAGT GCGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGCT CGAGTGGCTGCGTTCTTCATCGATGC
5a	CGTATCGCCTCCCTCGCGCCATCAGTAGTGT AGATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGAT CTGTCGGCTGCGTTCTTCATCGATGC
5b	CGTATCGCCTCCCTCGCGCCATCAGTAGTGT AGATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGCT CGAGTGGCTGCGTTCTTCATCGATGC
5c	CGTATCGCCTCCCTCGCGCCATCAGTAGTGT AGATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGTG ATGACGGCTGCGTTCTTCATCGATGC
6a	CGTATCGCCTCCCTCGCGCCATCAGTCGATC ACGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGCT CGAGTGGCTGCGTTCTTCATCGATGC
6b	CGTATCGCCTCCCTCGCGCCATCAGTCGATC ACGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGTG ATGACGGCTGCGTTCTTCATCGATGC
6c	CGTATCGCCTCCCTCGCGCCATCAGTCGATC ACGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTAT GTACAGGCTGCGTTCTTCATCGATGC
7a	CGTATCGCCTCCCTCGCGCCATCAGTCTATAC TATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCGTG ATGACGGCTGCGTTCTTCATCGATGC
7b	CGTATCGCCTCCCTCGCGCCATCAGTCTATAC TATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTAT GTACAGGCTGCGTTCTTCATCGATGC

8a	CGTATCGCCTCCCTCGCGCCATCAGTGACGT ATGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTAT GTACAGGCTGCGTTCTTCATCGATGC
8b	CGTATCGCCTCCCTCGCGCCATCAGTGACGT ATGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTCG ATATAGGCTGCGTTCTTCATCGATGC
8c	CGTATCGCCTCCCTCGCGCCATCAGTGACGT ATGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTCG CACGCGGCTGCGTTCTTCATCGATGC
9a	CGTATCGCCTCCCTCGCGCCATCAGTGTGAG TAGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTCG ATATAGGCTGCGTTCTTCATCGATGC
9b	CGTATCGCCTCCCTCGCGCCATCAGTGTGAG TAGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTCG CACGCGGCTGCGTTCTTCATCGATGC
9c	CGTATCGCCTCCCTCGCGCCATCAGTGTGAG TAGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTGC GTCACGGCTGCGTTCTTCATCGATGC
10a	CGTATCGCCTCCCTCGCGCCATCAGTGATAC GTCTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTCG CACGCGGCTGCGTTCTTCATCGATGC
10b	CGTATCGCCTCCCTCGCGCCATCAGTGATAC GTCTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTGC GTCACGGCTGCGTTCTTCATCGATGC
10c	CGTATCGCCTCCCTCGCGCCATCAGTGATAC GTCTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTGT GCGTCGGCTGCGTTCTTCATCGATGC
11a	CGTATCGCCTCCCTCGCGCCATCAGACATAC GCGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTGC GTCACGGCTGCGTTCTTCATCGATGC
11b	CGTATCGCCTCCCTCGCGCCATCAGACATAC GCGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTGT GCGTCGGCTGCGTTCTTCATCGATGC
11c	CGTATCGCCTCCCTCGCGCCATCAGACATAC GCGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTAGC ATACTGGCTGCGTTCTTCATCGATGC
12a	CGTATCGCCTCCCTCGCGCCATCAGACGCGA GTATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCTGT GCGTCGGCTGCGTTCTTCATCGATGC
12b	CGTATCGCCTCCCTCGCGCCATCAGACGCGA GTATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTAGC ATACTGGCTGCGTTCTTCATCGATGC
12c	CGTATCGCCTCCCTCGCGCCATCAGACGCGA GTATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTATAC ATGTGGCTGCGTTCTTCATCGATGC
13a	CGTATCGCCTCCCTCGCGCCATCAGACTACT ATGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTAGC ATACTGGCTGCGTTCTTCATCGATGC
13b	CGTATCGCCTCCCTCGCGCCATCAGACTACT ATGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTATAC ATGTGGCTGCGTTCTTCATCGATGC
13c	CGTATCGCCTCCCTCGCGCCATCAGACTACT ATGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTATCA CTCAGGCTGCGTTCTTCATCGATGC
14a	CGTATCGCCTCCCTCGCGCCATCAGACTGTA CAGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTATAC ATGTGGCTGCGTTCTTCATCGATGC
14b	CGTATCGCCTCCCTCGCGCCATCAGACTGTA CAGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTATCA CTCAGGCTGCGTTCTTCATCGATGC
14c	CGTATCGCCTCCCTCGCGCCATCAGACTGTA CAGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTATCT GATAGGCTGCGTTCTTCATCGATGC
15a	CGTATCGCCTCCCTCGCGCCATCAGAGACTA TACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTATCA CTCAGGCTGCGTTCTTCATCGATGC
15b	CGTATCGCCTCCCTCGCGCCATCAGAGACTA TACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTATCT GATAGGCTGCGTTCTTCATCGATGC
15c	CGTATCGCCTCCCTCGCGCCATCAGAGACTA TACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTCGT GACATGGCTGCGTTCTTCATCGATGC
16a	CGTATCGCCTCCCTCGCGCCATCAGAGCGTC GTCTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTATCT GATAGGCTGCGTTCTTCATCGATGC
16b	CGTATCGCCTCCCTCGCGCCATCAGAGCGTC GTCTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTCGT GACATGGCTGCGTTCTTCATCGATGC

16c	CGTATCGCCTCCCTCGCGCCATCAGAGCGTC GTCTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTCTG ATCGAGGCTGCGTTCTTCATCGATGC
17a	CGTATCGCCTCCCTCGCGCCATCAGAGTACG CTATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTCGT GACATGGCTGCGTTCTTCATCGATGC
17b	CGTATCGCCTCCCTCGCGCCATCAGAGTACG CTATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTCTG ATCGAGGCTGCGTTCTTCATCGATGC
17c	CGTATCGCCTCCCTCGCGCCATCAGAGTACG CTATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGAC ATCTCGGCTGCGTTCTTCATCGATGC
18a	CGTATCGCCTCCCTCGCGCCATCAGATAGAG TACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTCTG ATCGAGGCTGCGTTCTTCATCGATGC
18b	CGTATCGCCTCCCTCGCGCCATCAGATAGAG TACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGAC ATCTCGGCTGCGTTCTTCATCGATGC
18c	CGTATCGCCTCCCTCGCGCCATCAGATAGAG TACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGAG CTAGAGGCTGCGTTCTTCATCGATGC
19a	CGTATCGCCTCCCTCGCGCCATCAGCACGCT ACGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGAC ATCTCGGCTGCGTTCTTCATCGATGC
19b	CGTATCGCCTCCCTCGCGCCATCAGCACGCT ACGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGAG CTAGAGGCTGCGTTCTTCATCGATGC
19c	CGTATCGCCTCCCTCGCGCCATCAGCACGCT ACGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGAT AGAGCGGCTGCGTTCTTCATCGATGC
20a	CGTATCGCCTCCCTCGCGCCATCAGCAGTAG ACGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGAG CTAGAGGCTGCGTTCTTCATCGATGC
20b	CGTATCGCCTCCCTCGCGCCATCAGCAGTAG ACGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGAT AGAGCGGCTGCGTTCTTCATCGATGC
20c	CGTATCGCCTCCCTCGCGCCATCAGCAGTAG ACGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGCG TGTGCGGCTGCGTTCTTCATCGATGC
21a	CGTATCGCCTCCCTCGCGCCATCAGCGACGT GACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGAT AGAGCGGCTGCGTTCTTCATCGATGC
21b	CGTATCGCCTCCCTCGCGCCATCAGCGACGT GACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGCG TGTGCGGCTGCGTTCTTCATCGATGC
21c	CGTATCGCCTCCCTCGCGCCATCAGCGACGT GACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGCT AGTCAGGCTGCGTTCTTCATCGATGC
22a	CGTATCGCCTCCCTCGCGCCATCAGTACACA CACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGCG TGTGCGGCTGCGTTCTTCATCGATGC
22b	CGTATCGCCTCCCTCGCGCCATCAGTACACA CACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGCT AGTCAGGCTGCGTTCTTCATCGATGC
22c	CGTATCGCCTCCCTCGCGCCATCAGTACACA CACTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACGA CAGCTCGCTGCGTTCTTCATCGATGC
23a	CGTATCGCCTCCCTCGCGCCATCAGTACACG TGATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGTGCT AGTCAGGCTGCGTTCTTCATCGATGC
23b	CGTATCGCCTCCCTCGCGCCATCAGTACACG TGATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACGA CAGCTCGCTGCGTTCTTCATCGATGC
23c	CGTATCGCCTCCCTCGCGCCATCAGTACACG TGATGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACGT CTCATCGCTGCGTTCTTCATCGATGC
24a	CGTATCGCCTCCCTCGCGCCATCAGTACAGA TCGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACGA CAGCTCGCTGCGTTCTTCATCGATGC
24b	CGTATCGCCTCCCTCGCGCCATCAGTACAGA TCGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACGT CTCATCGCTGCGTTCTTCATCGATGC
24c	CGTATCGCCTCCCTCGCGCCATCAGTACAGA TCGTGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACTC ATCTACGCTGCGTTCTTCATCGATGC
25a	CGTATCGCCTCCCTCGCGCCATCAGTCATCG AGTCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACGT CTCATCGCTGCGTTCTTCATCGATGC

25b	CGTATCGCCTCCCTCGCGCCATCAGTCATCG AGTCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACTC ATCTACGCTGCGTTCTTCATCGATGC
25c	CGTATCGCCTCCCTCGCGCCATCAGTCATCG AGTCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACTC GCGCACGCTGCGTTCTTCATCGATGC
26a	CGTATCGCCTCCCTCGCGCCATCAGTCGAGC TCTCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACTC ATCTACGCTGCGTTCTTCATCGATGC
26b	CGTATCGCCTCCCTCGCGCCATCAGTCGAGC TCTCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACTC GCGCACGCTGCGTTCTTCATCGATGC
26c	CGTATCGCCTCCCTCGCGCCATCAGTCGAGC TCTCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGAG CGTCACGCTGCGTTCTTCATCGATGC
27a	CGTATCGCCTCCCTCGCGCCATCAGTCGCAG ACACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGACTC GCGCACGCTGCGTTCTTCATCGATGC
27b	CGTATCGCCTCCCTCGCGCCATCAGTCGCAG ACACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGAG CGTCACGCTGCGTTCTTCATCGATGC
27c	CGTATCGCCTCCCTCGCGCCATCAGTCGCAG ACACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGCG ACTAGCGCTGCGTTCTTCATCGATGC
28a	CGTATCGCCTCCCTCGCGCCATCAGTCTGTCT CGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGAG CGTCACGCTGCGTTCTTCATCGATGC
28b	CGTATCGCCTCCCTCGCGCCATCAGTCTGTCT CGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGCG ACTAGCGCTGCGTTCTTCATCGATGC
28c	CGTATCGCCTCCCTCGCGCCATCAGTCTGTCT CGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGTA GTGATCGCTGCGTTCTTCATCGATGC
29a	CGTATCGCCTCCCTCGCGCCATCAGTGAGTG ACGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGCG ACTAGCGCTGCGTTCTTCATCGATGC
29b	CGTATCGCCTCCCTCGCGCCATCAGTGAGTG ACGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGTA GTGATCGCTGCGTTCTTCATCGATGC
29c	CGTATCGCCTCCCTCGCGCCATCAGTGAGTG ACGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGTG ACACACGCTGCGTTCTTCATCGATGC
30a	CGTATCGCCTCCCTCGCGCCATCAGTGATGT GTACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGTA GTGATCGCTGCGTTCTTCATCGATGC
30b	CGTATCGCCTCCCTCGCGCCATCAGTGATGT GTACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGTG ACACACGCTGCGTTCTTCATCGATGC
30c	CGTATCGCCTCCCTCGCGCCATCAGTGATGT GTACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGTG TATGTCGCTGCGTTCTTCATCGATGC
31a	CGTATCGCCTCCCTCGCGCCATCAGTGCTAT AGACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGTG ACACACGCTGCGTTCTTCATCGATGC
31b	CGTATCGCCTCCCTCGCGCCATCAGTGCTAT AGACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGTG TATGTCGCTGCGTTCTTCATCGATGC
31c	CGTATCGCCTCCCTCGCGCCATCAGTGCTAT AGACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGATAG ATAGACGCTGCGTTCTTCATCGATGC
32a	CGTATCGCCTCCCTCGCGCCATCAGTGCTCG CTACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGAGTG TATGTCGCTGCGTTCTTCATCGATGC
32b	CGTATCGCCTCCCTCGCGCCATCAGTGCTCG CTACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGATAG ATAGACGCTGCGTTCTTCATCGATGC
32c	CGTATCGCCTCCCTCGCGCCATCAGTGCTCG CTACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGATAT AGTCGCGCTGCGTTCTTCATCGATGC
33a	CGTATCGCCTCCCTCGCGCCATCAGCGACAC TATCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGATAG ATAGACGCTGCGTTCTTCATCGATGC
33b	CGTATCGCCTCCCTCGCGCCATCAGCGACAC TATCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGATAT AGTCGCGCTGCGTTCTTCATCGATGC
33c	CGTATCGCCTCCCTCGCGCCATCAGCGACAC TATCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGATCTA CTGACGCTGCGTTCTTCATCGATGC

34a	CGTATCGCCTCCCTCGCGCCATCAGCGAGAC GCGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGATAT AGTCGCGCTGCGTTCTTCATCGATGC
34b	CGTATCGCCTCCCTCGCGCCATCAGCGAGAC GCGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGATCTA CTGACGCTGCGTTCTTCATCGATGC
34c	CGTATCGCCTCCCTCGCGCCATCAGCGAGAC GCGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCACG TAGATCGCTGCGTTCTTCATCGATGC
35a	CGTATCGCCTCCCTCGCGCCATCAGCGTATG CGACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGATCTA CTGACGCTGCGTTCTTCATCGATGC
35b	CGTATCGCCTCCCTCGCGCCATCAGCGTATG CGACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCACG TAGATCGCTGCGTTCTTCATCGATGC
35c	CGTATCGCCTCCCTCGCGCCATCAGCGTATG CGACGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCACG TGTCGCGCTGCGTTCTTCATCGATGC
36a	CGTATCGCCTCCCTCGCGCCATCAGCGTCGA TCTCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCACG TAGATCGCTGCGTTCTTCATCGATGC
36b	CGTATCGCCTCCCTCGCGCCATCAGCGTCGA TCTCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCACG TGTCGCGCTGCGTTCTTCATCGATGC
36c	CGTATCGCCTCCCTCGCGCCATCAGCGTCGA TCTCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCATA CTCTACGCTGCGTTCTTCATCGATGC
37a	CGTATCGCCTCCCTCGCGCCATCAGCTACGA CTGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCACG TGTCGCGCTGCGTTCTTCATCGATGC
37b	CGTATCGCCTCCCTCGCGCCATCAGCTACGA CTGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCATA CTCTACGCTGCGTTCTTCATCGATGC
37c	CGTATCGCCTCCCTCGCGCCATCAGCTACGA CTGCGGAAGTAAAAGTCGTAACAAGG	CTATGCGCCTTGCCAGCCCGCTCAGCAGA CGTCTGGCTGCGTTCTTCATCGATGC

Appendix B

Percent Detection of Species Presence between Subsamples

B.1 Table of all subsample comparisons used to determine percent detection of species presence

Sample	Subsample comparison	Detection of Species Present	Detection Percentage of Present Species	Average Species Detection Percentage per Sample
1	1a 1b	0.67	67	72.3
1	1b 1c	0.5	50	
1	1a 1c	1	100	
2	2a 2c	0.33	33	61.0
2	2b 2c	0.75	75	
2	2a 2c	0.75	75	
3	3a 3b	0.25	25	43.3
3	3b 3c	0.8	80	
3	3a 3c	0.25	25	
4	4a 4b	0.33	33	44.3
4	4b 4c	0.5	50	
4	4a 4c	0.5	50	
5	5a 5b	0.25	25	41.7
5	5b 5c	0.25	25	
5	5a 5c	0.75	75	
6	6a 6b	0.5	50	66.7
6	6b 6c	1	100	
6	6a 6c	0.5	50	
7	7a 7b	0.33	33	33.0
8	8a 8b	1	100	100.0
8	8b 8c	1	100	
8	8a 8c	1	100	
9	9a 9b	1	100	100.0
9	9b 9c	1	100	
9	9a 9c	1	100	
10	10a 10b	0.5	50	50.0
10	10b 10c	0.33	33	
10	10a 10c	0.67	67	
11	11a 11b	0.67	67	52.3
11	11b 11c	0.5	50	
11	11a 11c	0.4	40	
12	12a 12b	0.67	67	78.0
12	12b 12c	0.67	67	
12	12a 12c	1	100	
13	13a 13b	0.67	67	78.0
13	13b 13c	0.67	67	
13	13a 13c	1	100	

14	14a 14b	0.25	25	41.7
14	14b 14c	0.67	67	
14	14a 14c	0.33	33	
15	15a 15b	0.75	75	52.7
15	15b 15c	0.5	50	
15	15a 15c	0.33	33	
16	16a 16b	1	100	100.0
16	16b 16c	1	100	
16	16a 16c	1	100	
17	17a 17b	0.2	20	40.0
17	17b 17c	0.75	75	
17	17a 17c	0.25	25	
18	18a 18b	0.67	67	78.0
18	18b 18c	1	100	
18	18a 18c	0.67	67	
19	19a 19b	0.5	50	66.7
19	19b 19c	0.5	50	
19	19a 19c	1	100	
20	20a 20b	0.33	33	26.3
20	20b 20c	0.17	17	
20	20a 20c	0.29	29	
21	21a 21b	1	100	100.0
21	21b 21c	1	100	
21	21a 21c	1	100	
22	22a 22b	0	0	13.0
22	22b 22c	0.14	14	
22	22a 22c	0.25	25	
23	23a 23b	0.5	50	34.7
23	23b 23c	0.25	25	
23	23a 23c	0.29	29	
24	24a 24b	0.33	33	33.0
25	25a 25b	1	100	100.0
25	25b 25c	1	100	
25	25a 25c	1	100	
26	26a 26b	0.5	50	41.7
26	26b 26c	0.5	50	
26	26a 26c	0.25	25	
27	27a 27b	0.6	60	42.0
27	27b 27c	0.33	33	
27	27a 27c	0.33	33	
28	28a 28b	0.75	75	50.7
28	28b 28c	0.17	17	
28	28a 28c	0.6	60	
29	29a 29b	0.5	50	50.0
30	30a 30b	0.29	29	33.7
30	30b 30c	0.29	29	
30	30a 30c	0.43	43	
31	31a 31b	0.6	60	48.7
31	31b 31c	0.57	57	
31	31a 31c	0.29	29	

32	32a 32b	0.43	43	60.0
32	32b 32c	0.8	80	
32	32a 32c	0.57	57	
33	33a 33b	1	100	66.7
33	33b 33c	0.5	50	
33	33a 33c	0.5	50	
34	34a 34b	0.33	33	26.3
34	34b 34c	0.17	17	
34	34a 34c	0.29	29	
35	35a 35b	0.17	17	5.7
35	35b 35c	0	0	
35	35a 35c	0	0	
36	36a 36b	1	100	55.3
36	36b 36c	0.33	33	
36	36a 36c	0.33	33	
37	37a 37b	0.25	25	33.3
37	37b 37c	0.25	25	
37	37a 37c	0.5	50	

References

Aizen, M. A., Garibaldi, L. A., Cunningham, S. A., & Klein, A. M. (2008). Long-term global trends in crop yield and production reveal no current pollination shortage but increasing pollinator dependency. *Current Biology*, *18*, 1572-1575.

Aizen, M. A., & Harder, L. D. (2009). The global stock of domesticated honey bees is growing slower than agricultural demand for pollination. *Current Biology*, *19*(11), 1-4.

Al-Ghamdi, A., Adgaba, N., Getachew, A., & Tadesse, Y. (2014). New approach for determination of an optimum honeybee colony's carrying capacity based on productivity and nectar secretion potential of bee forage species. *Saudi Journal of Biological Sciences*, In Press.

Almeida-Muradian, L. B., Pamplona, L. C., Coimbra, S., & Barth, O. M. (2005). Chemical composition and botanical evaluation of dried bee pollen pellets. *Journal of Food Composition and Analysis*, *18*(1), 105-111.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol. 215*, 403-410.

Arita, L. H., Sakai, W. S., & Moniz, J. J. (1989). Pollen types collected by honey bees at three localities on the island of Hawaii part I: Floral source identification. *Hawaiian Entomological Society*, *29*, 125-132.

Baldwin, B.G. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the compositae. (1992). *Molecular Phylogenetics and Evolution*, 1 (1): 3-16.

Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S., and Donoghue, M.J. The ITS region of nuclear ribosomal DNA: A valuable source of evidence on Angiosperm phylogeny. (1995). *Annals of the Missouri Botanical Garden, 82 (2)*, 247-277.

Barth, O. M., Freitas, A. S., Oliveira, E. S., Silva, R. A., Maester, F. M., Andrella, R. R.S., & Cardozo, G. M. (2010). Evaluation of the botantical origin of commercial dry bee pollen load batches using pollen analysis: A proposal for technical standardization. *Anais da Academia Brasileira de Ciencias*, *82*(4), 893-902

Barth, O. M., Munhoz, M. C., & Luz, C. F.P. (2009). Botanical origin of Apis pollen loads using colour, weight, and pollen morphology data. *Acta Alimentaria*, *38*, 133-139.

Boch, R. 1982. Relative attractiveness of different pollens to honeybees when foraging in a flight room and when fed in the hive. *Journal of Apicultural Research, 21* (2), 104-106.

Bowie, M. H., Wratten, S. D., & White, A. J. (1995). Agronomy and phenology of "companion plants" of potential for enhancement of insect biological control. *New Zealand Journal of Crop and Horticultural Science*, *23*(4), 423-427.

Breeze, T.D., Bailey, A.P., Balcombe, K.G., & Potts, S.G. (2011). Pollination services in the UK: How important are honeybees? *Agriculture, Ecosystems and Environment*, *14*, 137-143.

Bruni, I., Galimberti, A., Caridi, L., Scaccabarozzi, D., De Mattia, F., Casiraghi, M., & Labra, M. (2015). A DNA barcoding approach to identify plant species in multiflower honey. *Food Chemistry, 170*, 308-315.

Camazine, S. (1993). The regulation of pollen foraging by honey bees: How foragers assess the colony's need for pollen. *Behavioral Ecology and Sociobiology*, *32*(4), 265-272.

Caporaso, J.G. et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods, 7*(5), 335-336.

Carreck, N., & Williams, I.H. (1997). Observations on two commercial flower mixtures as food sources for beneficial insects in the UK. *Journal of Agricultural Science*, *128*(4), 397-403.

Carvell, C., Meek, W. R., Pywell, R. F., Goulson, D., & Nowakowski, M. (2007). Comparing the efficacy of agri-environment schemes to enhance bumble bee abundance and diversity on arable field margins. *Journal of Applied Ecology*, *44*(1), 29-40.

Couvillon, M.J., Schürch, R., & Ratnieks, F.L.W. (2014). Dancing bees communicate a foraging preference for rural lands in high-level agri-environment schemes. *Current Biology, 24(11)*, 1212-1215.

Crane, E., Walker, P., & Day, R. (1984). *Directory of important world honey sources*. London: International Bee Research Association.

Di Pasquale, G., Salignon, M., Le Conte, Y., Belzunces, L.P., Decourtye, A., Kretzschmar, A., . . . Alaux, C. (2013). Influence of pollen nutrition on honey bee health: Do pollen quality and diversity matter? *PLoS ONE*, *8*(8).

Dimou, M., & Thrasyvoulou, A. (2007). A comparison of three methods for assessing the relative abundance of pollen resources collected by honey bee colonies. *Journal of Apicultural Research*, *46*(3), 144-148.

Dimou, M., Thrasyvoulou, A., & Tsirakoglou, V. (2006). Efficient use of pollen traps to determine the pollen flora used by honey bees. *Journal of Apicultural Research*, *45*(1), 42-46.

Dreller, C., & Tarpy, D.R. (2000). Perception of the pollen need by foragers in a honeybee colony. *Animal Behaviour*, *59*, 91-96.

Ekroos, J., Piha, M., & Tiainen, J. (2008). Role of organic and conventional field boundaries on boreal bumblebees and butterflies. Agriculture, Ecosystems & Environment: 124(3-4): 155-159.

Erdtman, G. (1943). *An Introduction to Pollen Analysis*. Massachusetts, USA: The Chronica Botanica Company.

European Commission. (2015). Agri-environment measures. Agriculture and Rural Development, European Commission. Retrieved from http://ec.europa.eu/agriculture/envir/measures/index_en.htm

Fewell, J.H., & Winston, M.L. (1995). Regulation of nectar collection in relation to honey storage levels by honey bees, Apis mellifera. *Behavioural Ecology*, *7*(3), 286-291.

Free, J. B. (1963). The Flower Constancy of Honeybees. *Journal of Animal Eology*, *32*(1), 119-131.

Galimberti, A., De Mattia, F., Bruni, I., Scaccabarozzi, D., Sandionigi, A., Barbuto, M., Casiraghi, M., & Labra, M. (2014). A DNA Barcoding approach to characterize pollen collected by honeybees. *PLoS ONE*.

García-García M.C., Ortiz, P.L., & Diez, M. (2004). Variations in the weights of pollen loads collected by Apis mellifera L. *Grana*, *43*, 183-192.

Genersch, E. (2010). American Foulbrood in honeybees and its causative agent, Paenibacillus larvae. *Journal of Invertebrate Pathology*, *103*, S10-S19.

Gonzalez-Martin, I., Hernandez-Hierro, J.M., Barros-Ferreiro, N., Marcos, C. C., & Garcia-Villanova, R.J. (2007). Use of NIRS technology with a remote reflectance fibre-optic probe for predicting major components in bee pollen. *Talanta*, *72*, 998-1003.

Gottardini, E., Rossi, S., Cristofolini, F., & Benedetti, L. (2007). Use of Fourier transform infrared (FT-IR) spectroscopy as a tool for pollen identification. *Aerobiologia*, *23*, 211-219.

Goulson, D. (2013). An overview of the environmental risks posed by neonicotinoid insecticides. *Journal of Applied Ecology*, *50*, 977-987.

Goulson, D., Cruise, J. L., Sparrow, K. R., Harris, A. J., Park, K. J., Tinsley, M. C., & Gilburn, A. S. (2007). Choosing rewarding flowers; perceptual limitations and innate preferences influence decision making in bumblebees and honeybees. *Behavioural Ecology and Sociology*, *61*, 1523-1529.

Henry, M., Frochen, M., Maillet-Mezeray, J., Breyne, E., Allier, F., Odoux, J.F., & Decourtye, A. (2012). Spatial autocorrelation in honeybee foraging activity reveals optimal focus scale for predicting agro-environmental scheme efficiency. *Ecological Modelling*, *225*, 103-114.

Hickman, J.M., & Wratten, S.D. (1996). Use of Phacelia tanacetifolia strips to enhance biological control of aphids by hoverfly larvae in cereal fields. *Journal of Economic Entomology*, *89*(4), 832-840.

Kearns, C. & Inouye, D.W. (1993). *Techniques for Pollination Biologists.* Colorado USA: University Press of Colorado.

Keller, I., Fluri, P., & Imdorf, A. (2005a). Pollen nutrition and colony development in honey bees, part I. *Bee World*, *86*(1), 3-10.

Keller, I., Fluri, P., & Imdorf, A. (2005b). Pollen nutrition and colony development in honey bees - part II. *Bee World*, *86*(2), 27-34.

Khansari, E., Zarre, S., Alizadeh, K., Attar, F., Aghabeig, F., & Salmaki, Y. (2012). Pollen morphology of Campanula (Campanulaceae) and allied genera in Iran with special focus on its systematic implication. *Flora- Morphology, Distribution, Functional Ecology of Plants*, *207*(3), 203-211.

Kleijn, D., Baquero, R. A., Clough, Y., Diaz, M., De Esteban, J., Fernandez, F., . . . Yela, J. L. (2006). Mixed biodiversity benefits of agri-environment schemes in five European countries. *Ecology Letters*, *9*, 243-254.

Kleijn, D., & Sutherland, W. J. (2003). How effective are European agri-environment schemes in conserving and promoting biodiversity? *Journal of Applied Ecology*, *40*(6), 947-969.

Klein, A.-M., Vaissiere, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society*, *274*, 303-313.

Kremen, C., Williams, N.M., Aizen, M.A., Gemmill-Herren, B., LeBuhn, G., Minckley, R., . . . Ricketts, T.H. (2007). Pollination and other ecosystem services produced by mobile organisms: A conceptual framework for the effects of land-use change. *Ecology Letters*, *10*(4), 299-314.

Kron, P. & Husband, B.C. (2012). Using flow cytometry to estimate pollen DNA content: improved methodology and applications. *Annals of Botany, 110(5)*: 1067-1078.

Kron, P., Kwok, A., & Husband, B. C. (2014). Flow cytometric analysis of pollen grains collected from individual bees provides information about pollen load composition and foraging behaviour. *Annals of Biology*, *113*, 191-197.

Laubertie, E. A., Wratten, S. D., & Hemptinne, J.-L. (2012). The contribution of potential beneficial insectary plant species to adult hoverfly (Diptera: Syrphidae) fitness. *Biological Control*, *61*, 1-6.

Leitch, I.J. & Bennett, M.D. (2007). Genome size and its uses: The impact of flow cytometry. In J. Dolezel, J. Greilhuber, & J. Suda (Eds.), *Flow cytometry with plant cells: Analysis of genes, chromosomes and genomes* (pp. 153-176). Weinheim: Wiley-VCH Verlag.

Levin, M.D. & Bohart, G.E. (1955). Selection of pollens by honey bees. *American Bee Journal, 95 (10)*: 392-397.

Marcos, J. V., Nava, R., Cristobal, G., Redondo, R., Escalante-Ramirez, B., Bueno, G., . . . Rodriguez, T. (2014). Automated pollen identification using microscopic imaging and texture analysis. *Micron*, *68*, 36-46.

Mitsumoto, K., Yabusaki, K., & Aoyagi, H. (2009). Classification of pollen species using autofluorescence image analysis. *Journal of Bioscience and Bioengineering*, *107*(1), 90-94.

Moore, P.D., Webb, J.A., & Collison, M.E. (1991). *Pollen Analysis*. Oxford, UK: Blackwell Scientific Publications.

Natural England. (2009). Agri-environment schemes in England 2009 (NE194). Natural England, Department for Environment, Food & Rural Affairs, UK. Retrieved from http://publications.naturalengland.org.uk/publication/46002?category=35001

Natural England. (2009). Farming with nature: Agri-environment schemes in action (NE223). Natural England, Department for Environment, Food & Rural Affairs, UK. Retrieved from http://publications.naturalengland.org.uk/publication/41005

Natural England. (2012). Evolution of Agri-environment schemes in England (NE373). Natural England, Department for Environment, Food & Rural Affairs, UK. Retrieved from http://publications.naturalengland.org.uk/publication/3567470?category=35001.

Natural England. (2013a). Entry level stewardship: Environmental Stewardship Handbook, Fourth Edition (NE349). Natural England, Department for Environment, Food & Rural Affairs, UK. Retrieved from http://publications.naturalengland.org.uk/publication/2798159?category=35001

Natural England. (2013b). Higher level stewardship: Environmental Stewardship Handbook, Fourth Edition (NE350). Natural England, Department for Environment, Food & Rural Affairs, UK. Retrieved from http://publications.naturalengland.org.uk/publication/2827091?category=35001

Newstrom-Lloyd, L., Scheele, F., Raine, I., Li, X., Gonzalez, M., & Roper, T. (n.d.). *Pollen pellet colour, purity and identification* [Pamphlet; PDF]. Retrieved from http://www.treesforbeesnz.org/__data/assets/pdf_file/0017/60254/pollen-pellet-colour-poster.pdf

Olsen, L., Hoopingardner, R., & Martin, E.C. (1979). Pollen Preference of honeybees sited on four cultivated crops. *Journal of Apicultural Research, 18* (3), 196-200.

Pernal, S. F., & Currie, R. W. (2002). Discrimination and preferences for pollen-based cues by foraging honeybees, Apis mellifera L. *Animal Behaviour*, *63*, 369-390.

Pettis, J. S., Lichtenberg, E. M., Andree, M., Stitzinger, J., Rose, R., & vanEngelsdorp, D. (2013). Crop pollination exposes honey bees to pesticides which alters their susceptibility to the gut pathogen Nosema ceranae. *PLOS ONE*, *8*(7), 1-9.

Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010a). Global pollinator declines: Trends, impacts and drivers. *Trends in Ecology and Evolution*, *25*(6), 345-353.

Potts, S. G., Roberts, S. P.M., Dean, R., Marris, G., Brown, M. A., Jones, R., . . . Settele, J. (2010b). Declines of managed honey bees and beekeepers in Europe. *Journal of Apiculltural Research*, *49*(1), 15-22.

Quince, C., Lanzen, A., Curtis, T.P., Davenport, R.J., Hall, N., Head, I.M., Read, F., & Sloan, W.T. Accurate determination of microbial diversity from 454 pyrosequencing data. (2009). *Nature Methods*, 6: 639-641.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rader, R., Howlett, B. G., Cunningham, S. A., Westcott, D. A., Newstrom-Lloyd, L. E., Walker, M. K., . . . Edwards, W. (2009). Alternative pollinator taxa are equally efficient but not as effective as the honeybee in a mass flowering crop. *Journal of Applied Ecology*, *46*, 1080-1087.

Ragubeer, C., Jurgens, A., & Johnson, S.D. (2013). The effect of nectar scent on honeybee (*Apis mellifera L.)* behaviour. *South African Journal of Botany*, *86*, 181.

Richardson, R.T., Lin, C.-H., Sponsler, D.B., Quijia, J.O., Goodell, K., & Johnson, R.M. (2015). Application of ITS2 metabarcoding to determine the provenance of pollen collected by honey bees in an agroecosystem. *Applications in Plant Sciences, 3(1)*: apps.1400066.

Salariato, D.L., Zuloaga, F.O., Cano, A., & Al-Shehbaz, I.A. Molecular phylogenetics of the tribe Eudemeae (Brassicaceae) and implications for its morphology and distribution. (2015). *Molecular Phylogenetics and Evolution*, 82A: 43-59.

Sammataro, D., & Avitabile, A. (2011). *The beekeeper's handbook* (4th ed.). Ithaca, NY: Comstock.

Sammataro, D., Gerson, U., & Needham, G. (2000). Parasitic mites of honey bees: Life history, implications, and impact. *Annual Review of Entomology*, *45*, 519-548.

Schmeider, R. & Edwards, R. (2011). Quality control and preprocessing of metagenomics datasets. *Bioinformatics, 27 (6)*: 863-864.

Schmidt, J.O. (1982). Pollen foraging preferences of honeybees. *The Southwestern Entomologist, 7* (4), 255-259.

Shen, M., Yang, X., Cox-Foster, D., & Cui, L. (2005). The role of varroa mites in infections of Kashmir bee virus (KBV) and deformed wing virus (DWV) in honey bees. *Virology*, *342*, 141-149.

Shendure, J. & Hanlee, J. (2008). Next-generation DNA sequencing. *Nature Biotechnology, 2*6, 1135-1145.

Stanley, R.G. & Linskens, H.F. (1974). *Pollen.* Berlin: Springer.

Stimec, J., Scott-Dupree, C. D., & McAndrews, J. H. (1997). Honey bee, *Apis mellifera*, pollen foraging in southern Ontario. *Canadian Field-Naturalist*, *111*, 454-456.

Tautz, J. (2008). *The buzz about bees: Biology of a superorganism* (D. C. Sandeman, Trans.). Springer.

Trees for Bees NZ. (2014, February 8). *Trees for Bees NZ Results for Protein Content of Pollen in selected plant species found to be visited by honey bees in New Zealand* [Fact sheet]. Retrieved May 12, 2015, from http://www.treesforbeesnz.org/__data/assets/pdf_file/0003/71328/tfb-results-for-protein-count.pdf

Westphal, C., Steffan-Dewenter, I., & Tscharntke, T. (2003). Mass flowering crops enhance pollinator densities at a landscape scale. *Ecology Letters 6* (11): 961-965.

Whittingham, M. J. (2007). Will agri-environment schemes deliver substantial biodiversity gain, and if not why not? *Journal of Applied Ecology*, *44*, 1-5.

Williams, C.S. (1997). Nectar secretion rates, standing crops and flower choice by bees on *Phacelia tanacetifolia*. *Journal of Apicultural Research*, *36*(1), 23-32.

Williams, I.H., & Christian, D.G. (1991). Observations on *Phacelia tanacetifolia* Bentham (Hydrophyllaceae) as a food plant for honey bees and bumble bees. *Journal of Apicultural Research*, *30*(1), 3-12.

Wilson, E. E., Sidhu, C. S., LeVan, K. E., & Holway, D. A. (2010). Pollen foraging behaviour of solitary Hawaiian bees revealed through molecular pollen analysis. *Molecular Ecology*, *19*, 4823-4829.

Winfree, R., Williams, N.M., Gaines, H., Ascher, J.S., & Kremen, C. (2008). Wild bee pollinators provide the majority of crop visitation across land-use gradients in New Jersey and Pennsylvania, USA. *Journal of Applied Ecology*, *45*(3), 793-802.

Woodcock, B. A., Edwards, M., Redhead, J., Meek, W. R., Nuttall, P., Falk, S., . . . Pywell, R. F. (2013).

Crop visitation by honeybees, bumblebees and solitary bees: Behavioural differences and diversity responses to landscape. *Agriculture, Ecosystems and Environment*, *171*, 1-8.

van der Zee, R., Pisa, L., Andonov, S., Brodschneider, R., Charriere, J.-D., Chlebo, R., . . . Wilkins, S. (2012). Managed honey bee colony losses in Canada, China, Europe, Israel and Turkey, for the winters of 2008-9 and 2009-10. *Journal of Apicultural Research*, *51*(1), 100-114.

vanEngelsdorp, D., Caron, D., Hayes, J., Underwood, R., Henson, M., Rennich, K., . . . Pettis, J. (2012). A national survey of managed honey bee 2010-11 winter colony losses in the USA: results from the Bee Informed Partnership. *Journal of Apicultural Research*, *51*(1), 115-124.

vanEngelsdorp, D., & Meixner, M. D. (2010). A historical review of managed honey bee populations in Europe and the United States and factors that may affect them. *Journal of Invertebrate Pathology*, *103*, S80-S95.

Vilsack, T. & McCarthy, G. National strategy to promote the health of honey bees and other pollinators. Report issued by the White House and the Pollinator Health Task Force on 19 May 2015.